1

A Quantitative Model of Expert Transcription Typing

Bonnie E. John[1]

8 March 1993
CMU-CS-93-120

School of Computer Science
Carnegie Mellon University
Pittsburgh, PA, 15213-3890

Abstract

The Model of Expert Transcription Typing (METT) is a quantitative model built within the architecture of the Model Human Processor (MHP, Card, Moran & Newell, 1983). As such, it can be used to make quantitative predictions of performance on typing tasks, can be integrated with other MHP-based models of performance (e.g., choice reaction time), and can be extended to other typing-like domains (e.g., a data-entry task). In this paper, I apply the METT to transcription typing tasks that display robust behavioral phenomena (Salthouse, 1986) and evaluate its performance relative to these tasks. I briefly discuss integration with other MHP-based models and extension of the METT to predict performance in a complex, real-world task.

This research was supported in part by the Office of Naval Research, Perceptual Science Program, Contract Number N00014-87-K-0432 and the Cognitive Science Program, Contract Number N00014-89-J-1975N158. The views and conclusions contained in this document are those of the author and should not be interpreted as representing the official policies, either expressed or implied, of the Office of Naval Research or the U. S. Government.

[1]School of Computer Science and Department of Psychology

Keywords: GOMS, Cognitive Simulation, Human Information Processing, Model Human Processor, Typing

MOTIVATION

Two distinct but complementary research agendas motivate our study of transcription typing: constructing unified theories of cognition and the formation of engineering models of human behavior.

Allen Newell's position on cumulative, unified theory building as the appropriate mode for research (as opposed to isolated hypothesis-testing) has a well articulated history (Newell, 1973; Card, Moran, & Newell, 1983; Newell & Card, 1985, 1986; Newell, 1990). His suggestions for research directions in psychology include building complete processing models, taking particular tasks and analyzing all the psychological phenomena involved in them, and applying a model of a single phenomenon to lots of diverse situations where that phenomenon arises (Newell, 1973). The definition of theory he uses is endemic to all science: a "body of explicit knowledge, from which answers can be obtained to questions by inquiries" (Newell, 1990, p.13), where the answers can be explanations, predictions, or prescriptions for control . He expanded this definition to include the concepts of completeness and deliberate approximation. Newell also believed in benefits to psychology of attempting to apply psychological theory in a pragmatic domain, e.g., designing human computer interfaces. Such investigations, he said, "reveal new phenomena and set forth clearly what it is that needs explanation" (Card et al., 1983, p. 14). Newell comes to these notions of completeness, prediction, approximation, and application, from his vision of what will most benefit the scientific study of mind. My own position comes from the opposite direction, from a vision of what type of psychological theory will most benefit the design of computer systems. I have argued that models of human behavior useful in engineering design must parallel models in other engineering disciplines: they must make quantitative, a priori predictions, they must cover complete tasks in an integrated way, they must be deliberately approximate, and they must be usable by engineers not trained in the interpretation of psychological data (John, 1988).

A similar set of parallel motivations combine to focus our attention on the domain of transcription typing. From the side of pure psychology, several researchers have argued that transcription typing is a particularly good activity for the study of human skilled behavior (e.g., Rumelhart & Norman, 1982; Salthouse 1986). Their reasons are both scientific (e.g., typing involves a complex interaction of perception, cognition and motor activities) and pragmatic (e.g., there is an easily available population of individuals at different levels of skill). From the side of engineering design, typing is the most common form of human-to-computer communication today, hence an understanding of typing is necessary for a complete model of computer use. In addition, the release of a paper by Timothy Salthouse (1986) laying out 29 independently determined, robust phenomena of transcription typing, was no small motivation to construct a model of that activity.

These varied motivations combine to produce the model of expert transcription typing (METT) I present in this paper. In distilling the typing literature, Salthouse sought to define "the criteria by which alternative models in this area may be evaluated" (Salthouse, 1986, p. 304) and I present our alternative for such evaluation. Motivations aside, this model is the first we know of to run Salthouse's gauntlet of empirical phenomena.

In the remainder of this paper, I will first review Salthouse's 1986 paper and the state of theories of typing. I will then present the theoretical framework behind the METT and the that define the 29 phenomena included in Salthouse's gauntlet and evaluate its predictions. I will discuss how the METT faired on its run, its strengths and weaknesses, and the ways in which future research could strengthen its performance. Lastly, I will discuss the METT's contribution to the development of a unified theory of cognition and to an engineering model of human performance applicable to computer design.

THE "SALTHOUSE 29" AND THEORIES OF TYPING

A gauntlet of robust empirical pnenomena stands ready to test any theory of transcription typing performance. Timothy Salthouse distilled the literature detailing the phenomena associated with typing and presented 29 robust phenomena. As mentioned above, his stated purpose was to define "the criteria by which alternative models in this area may be evaluated" (Salthouse, 1986, p. 304).

Although several models of typing exist, none to date have attempted a quantitative run of the Salthouse gauntlet. As an indication of the state of theoretical development concerning typing, Salthouse said,

> Ideally, this should be done for several alternative models simultaneously to provide a basis for comparative evaluation, but this is not yet practical, either because competing models have not yet been specified in sufficient detail to derive explanations or because the other models were intended to apply to only a limited set of typing processes. (p. 304)

For instance, Salthouse proposed a model of transcription typing able to provide qualitative explanations for the phenomena he identified, but not quantitative predictions. It was, "derived from ideas introduced by earlier theorists (e.g. Cooper, 1983; Logan, 1983; Rumelhart & Norman, 1982; Shaffer, 1973, 1975a [referenced in this paper as Shaffer, 1975], 1976; Shaffer & Hardwick, 1970; Thomas & Jones, 1970)...[to be used] as a heuristic device to help organize [his] review of the empirical literature" (p. 303). His model is based on four processing components: input, which converts text into chunks; parsing, which decomposes chunks into ordinal strings of characters; translation, which converts characters into movement specifications; and execution, which implements movements in a ballistic fashion.

A model of typing presented by Rumelhart and Norman (1982) is "based upon an Activation-Trigger-Schema system in which a hierarchical structure of schemata directs the selection of the letters to be typed and, then, controls the hand and finger movements by a cooperative, relaxation algorithm" (p. 8). It is embodied in a computer simulation and provides detailed predictions about the movement of fingers, the relative response times for letters in different contexts, and several types of errors. However, Rumelhart and Norman explicitly state that their model does not cover

> ...the mechanisms used by inexperienced typists, nor the mechanisms involved in learning...the mechanisms involved in perception or the encoding of strings to be typed, nor in monitoring the accuracy of the typing...the deterioration of typing rate that occurs as the text is modified from normal prose to non-language or random letters...[or] non-alphabetical keys. (p. 6)

Rumelhart and Norman also do not attempt to make zero-parameter quantitative predictions of typing performances. Their predictions are made in terms of "arbitrary model units" (Figure 8, p. 23), which require the introduction of at least two parameters to convert to predictions of absolute performance times.

Sternberg, Knoll, and Wright (1978) present a subprogram-retrieval model. This model, again, is intended to apply to a subset of typing phenomena, namely the motoric aspects of typing. The model is based on research using a discontinuous typing task where a word, or string of characters, is presented well before the subject must type the string. Thus, "the perception of the material to be typed has presumably all occurred early in the trial, and the subject has plenty of time to rehearse or prepare in other ways for what he or she has to type" (p. 4). This model qualitatively accounts for several phenomena associated with discontinuous typing including the dependence of latency and interkeystroke interval on

string-length, and a serial position effect on interkeystroke interval. However, it does not account for phenomena arising from the interplay of perception, cognition, and motor behavior displayed in continuous typing.

In contrast to these models of typing, the METT is a quantitative, integrated model of perception, cognition and motor behavior in typing. Since the METT is both quantitative and integrated it is prepared to make quantitative predictions about more aspects of typing performance than previous models.

THE MODEL OF EXPERT TRANSCRIPTION TYPING

The METT is an information processing model of typing behavior. It is one of a family of models within the framework of the Model Human Processor (MHP) described by Card et al. in *The Psychology of Human-Computer Interaction* (1983). The MHP was an attempt to create a unified, integrated, model of human behavior that would allow quantitative predictions of behavior, especially in the domain of human-computer interaction (HCI). The MHP is specified in terms of processors, memories and a few quantitative parameters of each. *The Psychology of Human-Computer Interaction* , and a host of subsequent papers, went on to apply this basic theoretical framework to many areas of HCI: using text editors, graphic editors, and operating systems (Card et al., 1983), learning command languages (Bovair, Kieras, & Polson, 1990; Kieras & Polson, 1985; Polson & Kieras, 1985; Ziegler, Hoppe, & Fahnrich, 1986), using command abbreviations (John, Rosenbloom, & Newell, 1985; John & Newell, 1987, 1990), error behavior with spreadsheets and databases (Lerch, Mantei, & Olson, 1989;Smelcer, 1989), performance with different workstations for telephone operators (Gray, John, & Atwood, 1991, 1992; Gray, John, Stuart, Lawrence, & Atwood, 1990), among others.

The MHP provides the theoretical framework of the METT. An understanding of the elements of the MHP is necessary to assimilate the METT and its application to the typing tasks examined here. Therefore, a brief summary of these elements will be presented; greater detail can be found in the original work.

The Model Human Processor

There are three separate processors in the MHP: the perceptual processor, the cognitive processor and the motor processor (see Figure 1). These processors work serially within themselves, but in parallel with each other, subject to serial limitations imposed by data flow requirements (e.g., the cognitive processor may need information from the perceptual processor before it can proceed). The perceptual processor receives information from the outside world and deposits it into sensory memory, the most important of which are the Visual Image Store (VIS) and the Auditory Image Store (AIS), where it is held while being symbolically encoded. The Working Memory (WM) receives the symbolically encoded information and the cognitive processor uses that information, together with previously stored information from Long Term Memory (LTM), to make decisions about what to do. The cognitive processor deposits information into WM that initiates action by the motor processor. The motor processor acts on the outside world (e.g., presses keys, responds verbally, etc.).

The processors and memories are described by a few parameters. Card et al. (1983) derived estimates of these parameters through the distillation of relevant research in cognitive psychology.

The parameter describing each processor is its cycle time, τ. The perceptual processor cycle

time, τ_P ranges from 50 to 200 msec, with a typical value of 100 msec. The cognitive processor cycle time, τ_C, ranges from 25 to 170 msec, with a typical value of 70 msec. The motor processor cycle time, τ_M, ranges from 30 to 100 msec, with a typical value of 70 msec.

The processor cycle times are given as fairly wide ranges. They were set to cover the range of times found in the literature without establishing any substructure to the processors. That is, the processors are not specified to the level at which their substructure can be examined, so all operations performed by the processors are assumed to be elementary and the differences in operators are evident only in the wide range of values found when the cycle times are estimated with different tasks.

The parameters describing each memory are its storage capacity in items, μ, the decay time of an item, δ, and the main code type (physical, acoustic, visual, semantic), κ. Values for these parameters for each of the memories are shown in Figure 1. Only the capacity of WM is particularly relevant to the METT. The pure storage capacity of WM, μ_{WM}, ranges from 2.5 to 4.1 chunks, with a typical value of 3 chunks. If you allow unpacking of chunks from LTM, then the effective storage capacity of WM, μ_{WM^*}, ranges from 5 to 9 chunks, with a typical value of 7 chunks.

The processors and memories of the MHP work together under the control of ten principles of operation (Figure 1). These principles cover many aspects of the operations of the three processors, ranging from statements about the cycle times of the processors to a general statement that problem solving takes place in a problem space. They are derived from psychological evidence (e.g., a preponderance of data confirming Fitts's Law) and from a specific model of the psychological world (e.g., P8: the rationality principle). The principles of operation, like the wide ranges for processor cycle times, are an approximation. If the processors were computationally fully specified, then these effects would not be a list of principles, but a direct consequence of the operation of the processors. Although Card et al. did not produce a fully integrated model, the approximation provided by the principles of operation is sufficient for many tasks and makes for a much simpler model.

Recent work by John and Newell (1990) in the domain of stimulus-response compatibility, was specifically intended to produce estimates of operator durations that would be transferable to other tasks. Such estimates allow zero-parameter, quantitative predictions of performance to be made about tasks that use those operators. Three of the operators estimated are relevant to the task of transcription typing (Table 1). The first is a re-estimate of the cognitive cycle time, with experiments designed to reveal a value for a minimal cognitive operator. The remaining two are composed of several actions of the MHP processors and are operationally defined: (1) a perceptual operator perceives a written word and encodes it into an ordered list of letters that is the spelling, (2) a motor operator finds and hits a key on a keyboard (assuming an unskilled typist of about 30 wpm). These operators and their duration estimates are transferred directly to the METT (see ASSUMPTION 6), and are used to make quantitative predictions of performance.

The Assumptions of the METT

A few typing-specific assumptions are added to the MHP's basic three-processor structure to form the METT. These assumptions were arrived at through an examination of

transcription typing tasks, particularly the information flow between processors, and represent one set of possible assumptions.

ASSUMPTION 1: BASIC METHOD. The basic method is that a person perceives something (word, syllable, letter) and encodes it into an ordered list of letters (the spelling) with a perceptual operator. If it is a word or syllable, the spelling of that unit is obtained from memory with a cognitive operator, the first character is initiated with a cognitive operator, and then executed with a motor operator. The second character is then initiated and executed, and so on. If a letter is perceived, then the character is initiated immediately following the perception, and executed. The method for the word or syllable process is as follows (the letter process skips L3).[2]

```
                                                            Operator Type
BEGIN                                                  L 1
?Chunk ← Get-Chunk("To-Be-Typed-Copy")                 L 2  Perceptual
?Spelling ← Get-Spelling(?Chunk)                       L 3  Cognitive
?Letter ← Initiate-Letter(First-Letter[?Spelling])     L 4  Cognitive
Execute-Letter(?Letter)                                L 5  Motor
REPEAT BEGIN                                           L 6
    ?Letter ← Initiate-Letter(Next-Letter[?Spelling])  L 7  Cognitive
    Execute-Letter(?Letter)                            L 8  Motor
    END  L 9
 UNTIL Done?(?Spelling)                                L10
 IF-SUCCEEDED Done?(?Spelling)                         L11  Cognitive
   THEN END                                            L12
END                                                    L13
```

ASSUMPTION 2: SERIAL/PARALLEL PROCESSING. The following constraints on the parallel operation of the three processors are assumed for typing tasks.

a. PERCEPTION/COGNITION INTERACTION Perception has to be complete before getting the spelling or initiation of a character can begin.
b. SAME-HAND CONSTRAINT. A character on the same hand cannot be initiated with a cognitive operator until the motor processor execution of the previous character is complete.
c. PERCEPTION/WM LIMITATION INTERACTION. The perceptual processor cannot perceive the next piece of information unless there is room in WM for that information (see assumption 3 for the implication of this).

[2]The following algorithm is expressed in an informal language called ARTLess, introduced by Rosenbloom (1983; Laird, Rosenbloom & Newell, 1986) and subsequently refined by John, Rosenbloom and Newell (1985; John & Newell, 1987, 1990). Although the gist of the algorithms can be understood by readers familiar with procedural programming languages through a "natural-language" reading, the following features of the language may clarify some of the details. 1) Artless has seven constructs: operators(x), constants ("x"), variables(?x), assignments(←), blocks (BEGIN...x...END), branches (IF x THEN y), and loops (REPEAT x UNTIL y IF-SUCCEEDED y). 2) Each operator has an identifiable duration (i.e., a "cost") that is dependent on the processor performing the operator; manipulating constructs other than through operators is without cost. 3) ARTless also has a data structure called an ordered list of letters. An ordered list of letters has two functions defined to act upon it: get the first element of the ordered list (First-Letter) and get the next element of the ordered list (Next-Letter). If the list is empty, First-Letter will fail. Next-Letter returns the successor to the member of the list last returned. If the last-returned member of the list is the last member of the list, then Next-Letter will fail. This data structure is used to represent the spelling of words or syllables (i.e., chunks). When a chunk is being typed out, once the spelling of the chunk is gotten, the individual letters come off the ordered list with First-Letter and Next-Letter without cost. 4) Tests in conditionals must be exhaustive and explicit; there is no IF-THEN-ELSE construct. 5) Tests to exit a loop cost only when they succeed.

ASSUMPTION 3: WM LIMITATION. In normal transcription typing, the perceptual processor stays three chunks ahead of the cognitive processor. The chunk is usually a word, but it can be a syllable or a character if words are not available in the specific typing task. This three-chunk limitation is the typical capacity of working memory proposed in the MHP. These chunks may access long term memory (LTM) for the spelling of one chunk at a time, extending the effective capacity of working memory to, on average, seven chunks (assuming an average word length of five letters, the seven chunks would be the five letters of the first word, plus the two next words).

ASSUMPTION 4: PERCEPTUAL CHUNKS. The perceptual processor perceives at the most meaningful level available at or below the word level. For example, if words are present, they are perceived and encoded. If the view of whole words is restricted or if there are no words present (as when typing random letters), the perceptual processor perceives syllables. If syllables are not visible because of restricted view or random characters, then the perceptual processor perceives single characters.

ASSUMPTION 5: COGNITION/MOTOR INTERACTION. Once a character is initiated with a cognitive operator, the motor operator that executes that character cannot be stopped.

ASSUMPTION 6: OPERATOR SIMILARITY. Across all domains to which the MHP has been applied, similar operations involving similar perceptual, cognitive, and motor operators take similar amounts of time. Thus cognitive and perceptual operators and their duration estimates are transferred directly from the stimulus-response compatibility work of John and Newell (Table 1) to the METT in the next two assumptions. (The motor operator duration in Table 1, derived from inexperienced typists, serves as an extreme upper bound for expert-typist motor operators, and is used in the next section to estimate more realistic expert-typist motor operator durations.)

a. PERCEPTUAL OPERATOR DURATION. The time to perform a perceptual operator (perceive a visual stimulus) is 340 msec. A simplifying assumption is that this time is constant even if the thing to be perceived is a word, a syllable, or a character.
b. COGNITIVE OPERATOR DURATION. The time to perform a cognitive operator is 50 msec.

ASSUMPTION 7: MOTOR OPERATOR DURATION AND INTERACTION WITH SKILL. As a simplifying assumption, practice in typing decreases the motor operator time only; the estimates of the perceptual and cognitive operators remain constant. There are undoubtedly individual differences in the perceptual and cognitive processes, but given the amount of practice an adult typist has had perceiving words (as a part of reading) and in the cognitive operations involved in spelling words (as a part of writing), we assume that these operators remain constant relative to the more newly acquired motor operators of typing.

It is interesting to note the similarities between the model proposed by Salthouse and the METT. His input component corresponds to the Get-Chunk perceptual operator, his parsing component corresponds to the Get-Spelling cognitive operator, his translation component corresponds to the Initiate-Letter cognitive operator, and his execution component corresponds to the Execute-Letter motor operator. The METT comes directly from the structure of the MHP; it's similarity to a model stemming from years of

experience with the empirical data of typing lends credence to the MHP architecture.

Estimating the Motor Operator

The METT says that the duration of the motor operator decreases with typing skill. In order to make predictions of different typing tasks with differently skilled typists, we need to provide estimates for the motor operator associated with different typing speeds. This derivation of the motor operator will also serve as a an introductory example of how to use the METT.

Consider a typist typing a sentence from a standard typing test,[3] the *example sentence*, which will be used in analyses throughout this work):

> "One reason is quite obvious; you can get in and get out without waiting for the elevator."

The first three words are perceived with three perceptual operators, the spelling of the first word is retrieved from LTM with a cognitive operator, and the letters of the word, and the space following it, are initiated and executed in turn. As soon as the space has been initiated, the chunk is out of working memory, making room for the next word, so the perceptual processor perceives the next word. The processes continue until the entire sentence is typed.

The parallel operation and sequential dependencies of the three processors make the processes of typing difficult to analyze and talk about. Fortunately, there is an analysis technique borrowed from engineering project management that allows easy analysis of parallel resources (the three processors) working with sequential dependencies (outlined by the typing-specific assumptions). The technique is called critical path analysis. The particular version of this technique used for the analyses in this paper is embodied in a project management software package for the Apple Macintosh family of computers, MacProject.[4]

In a critical path analysis, each *subtask*, or *operator*, necessary to accomplish a total task is represented as a box with a duration (e.g., Get-chunk would be a subtask of transcription typing that would have a duration of 340 msec.). Dependencies between subtasks are represented by lines connecting the boxes (e.g., the cognitive operator, Get-Spelling, could not be started until the perceptual operator, Get-chunk, is completed, so a line would be drawn from Get-chunk to Get-Spelling). The representation of a full task is called a *schedule chart.*. The critical path is that set of subtasks that must be accomplished within their stated duration, or the entire task will take longer than anticipated.

Critical path was used to make estimates of the motor operator for typist of different speeds. By definition, a 60 gwpm typist would be able to type the 89 characters of the example sentence in 17,800 msec. The schedule chart for typing the example sentence (Figure 2) is drawn up with all perceptual operators along the top row, with the boxes labeled with the word the typist sees in the to-be-typed-copy (i.e., the example sentence). The cognitive operators are along the center row, labeled with the result of the Get-Spelling operator (when the content of the box is the word previously perceived), or the character to be initiated with an Initiate-Letter operator. The motor operators are along the bottom row labeled with the character they are typing with the Execute-Letter operator. The duration of

[3]Typing test obtained from the Carnegie Mellon Univeristy Personnel Office, 1987.

[4]Apple, Macintosh and MacProject are trademarks of Apple Computer, Inc.

all perceptual operators is set to 340 msec, the duration of all cognitive operators set to 50 msec, and the duration of all motor operators set to an initial guess of 230 msec (the motor operator found in the S-R compatibility research). The critical path is calculated and the operators on the path are counted: 1 perceptual, 38 cognitive and 90 motor operators. Given the respective durations of these operators, the total time to complete this sentence would be

$$1 * 340 \text{ msec} + 38 * 50 \text{ msec} + 90 * 230 \text{ msec} = 22{,}940 \text{ msec}$$

This is too long for a 60 gwpm typist, so the initial guess for the motor operator duration is too high, not surprisingly because that estimate was for a 30 gwpm typist. Iterating through this process yields a schedule chart with the same critical path, but the motor operator estimate giving the total time closest to 17,800 msec is 170 msec.

As well as getting numerical estimates from the critical path analyses, qualitative information about the roles of the three processors is obtained through inspection of the schedule chart. For instance, the perceptual operators are never on the critical path once the initial words have been perceived. This implies that the perceptual processes are not the limiting factors in the typing task. On the other hand, for a 60 gwpm typist, all the motor operators are on the critical path, indicating that a speed up of the motor operator will greatly affect the total time. At the other end of the speed spectrum, 160 gwpm, the critical path (Figure 3) looks quite different, with the cognitive operators determining the critical path and the motor operators playing a much less important role. This is because the duration of the motor operator is now shorter than that of the cognitive operator. This implies a theoretical maximum for typing, 180 gwpm, when the motor operator goes to zero, given the simplifying assumption that all the speed-up with skill comes from a decrease in the motor operator (Assumption 6c)

Repeated application of this analysis process yields a chart (Figure 4) of estimates of the motor operator vs. the gross speed of the expert typists used in the studies reviewed by Salthouse (1986). The motor operator estimates are rounded to the tens digit when used to analyze typing tasks. Armed with the METT, the critical path analysis technique, and estimates for the motor operators of different speed typists, the 29 phenomena presented by Salthouse can now be analyzed in detail.

EXPLAINING THE SALTHOUSE 29 WITH THE METT

Salthouse (1986)[5] separated his 29 phenomena into four categories: basic phenomena, units of typing, errors, and skill effects. This section follows the categorization and enumeration of the phenomena presented by Salthouse, and uses them as a supply of phenomena and data for the METT to explain. Each phenomenon has been rated as to whether it is parametric (the results are given as a function of parameter, e.g., typing speed, preview window size), quantitative (an average numerical value is reported, but no significant pattern of results with another parameter), or qualitative (the phenomenon is a relative comparison rather than a numeric value).

For each phenomenon in turn, the METT is used to simulate a typist's performance on the tasks used to define that phenomenon, producing predictions of performance. Corresponding to the ratings of the phenomena themselves, the METT account of each phenomenon is rated as parametric, quantitative, qualitative, not covered, or contradictory.

---

[5]Unless otherwise noted, in the remainder of this section all citations to Salthouse refer to the Salthouse, 1986 paper.

The goal of the METT is to make quantitative, zero-parameter predictions of performance, therefore, a measure of the error between the prediction and the observed behavior is preferable than, say, the amount of variance explained by the model. Also, the absolute magnitude of the error is not as important as the proportion of error in a prediction. Therefore, I use percent error to evaluate predictions of the METT, defined by:

$$\text{Percent Error} = \frac{(\text{Observed Performance} - \text{Predicted Performance})}{\text{Observed Performance}} \times 100\%.$$

Consonant with the ubiquitous engineering 80/20 rule (i.e., 80% of results come from 20% of the effort, but it takes 80% more effort to accomplish that last 20% of results), we count METT predictions as being parametric if across the range of the parameter the average absolute percent error is less than 20%. Likewise, a METT prediction is counted as quantitative if it is within 20% of the observed average behavior. Finally, if the METT makes a quantitative prediction that is in the same direction as the observed behavior, but is outside the 20% error criterion, we count that prediction as a qualitative success. A METT prediction can be as precise as the available data in this hierarchy, but not more so. For example, predictions could be qualitative for a quantitative phenomenon, but not parametric. At the end, the success of the METT will be evaluated, compared to the available phenomena as characterized by Salthouse.

Before examining Salthouse's gauntlet of phenomena there are two limitations on the scope the METT that should be emphasized: the METT distinguishes only between the two hands, not between different fingers, and it is restricted to expert performance (typists with speeds over 60 wpm). Therefore, we should not expect the METT to make predictions about phenomena that require a distinction between fingers (Phenomena 8, 11, and 23). Although the METT also cannot make predictions about phenomena that involve the full transition from novice to expert, it can make predictions for skill effects within the expert range (above 60 gwpm). While we would not expect the effects of skill to be as pronounced within the expert range, we will examine these predictions of the METT for the phenomena that report results in that range (Phenomenon 23-29).

Basic Typing Phenomena

The first twelve phenomena are basic typing phenomena. These phenomena involve the speed of typing relative to other tasks (phenomena 1-3), how degradation of the text away from normal prose affects the rate of typing (phenomena 4-6), patterns of interkey intervals (phenomena 7-11), and the effects of a concurrent task (phenomenon 12).

Phenomenon 1: Typing is Faster than Choice Reaction Time

> People can type very quickly, with interkey intervals averaging only a fraction of the typical choice reaction time...[For an average typists of 63 wpm] the median interkey interval in normal transcription typing was 177 ms. whereas the median interkey interval for the same individuals in a serial two-alternative choice reaction time task was 560 ms. *(Salthouse, 1986, p. 304)*

From the serial/parallel assumption and the same-hand constraint assumption, the interkey interval during transcription typing is either one motor operator for digrams that alternate hands, or one cognitive operator plus one motor operator for digrams on the same hand. Thus, the average interkey interval for a 60 gwpm typist (motor operator = 170 msec), assuming an equal number of same- and alternate-hand sequences, would be 195 msec, 10.2% away from the 177 msec. observed for a 63 wpm typist.

Typical times quoted for a two-choice reaction time task are about 300-400 msec (Luce, 1986),which, although substantially lower than Salthouse's 560 msec, still supports the observed phenomenon. Salthouse's two-choice reaction time task was somewhat non-standard:

> Stimuli were uppercase and lowercase versions of the letters L and R, and responses were presses of the leftmost and rightmost keys on the lowest row on the keyboard, Z (for l and L) and / (for r and R). Subjects were instructed to respond as rapidly and as accurately as possible. Each keystroke caused the immediate display of the next stimulus until a total of 50 randomly arranged stimuli had been presented. (Salthouse, 1984a, p. 350)

Using the model of immediate behavior presented in previous stimulus-response compatibility work (John, et. al. 1985; John & Newell, 1987, 1990), we can write a minimal algorithm for this particular two-choice reaction time task.

```
                                                    Operator type
BEGIN
?Letter ← Get-Stimulus("Letter")                    Perceptual
IF-SUCCEEDED Is-R?(?Letter)                         Cognitive
  THEN BEGIN
        Initiate("Right-Hand")                      Cognitive
        Execute("Right-Hand")                       Motor
        END
  ELSE IF-SUCCEEDED Is-L?(?Letter)                  Cognitive
          THEN BEGIN
                Initiate("Left-Hand")               Cognitive
                Execute("Left-Hand")                Motor
                END
END
```

In this algorithm, there are one perceptual operator (340 msec), two or three cognitive operators (for an average of 2.5 * 50 msec or 125 msec), and one motor operator (170 msec), for a total of 635 msec (13.4% above the observed 560 msec).
The METT explanation for this phenomenon is evident from the algorithms used to simulate the tasks. The typing algorithm allows overlap of the perceptual, cognitive and motor processes, and thus the average interkeystroke time (assuming equal same- and alternate-hand keystrokes) is the average of a single motor operator (interkey interval for alternate-hand keystrokes) and a motor operator plus a cognitive operator (interkey interval for same-hand keystrokes). In the CRT task, the algorithm is totally serial; the stimulus must be perceived, then cognitively processed, then the response is made. No overlap of processes means that all of the operators are on the critical path and contribute to the longer interkeystroke interval.

This is a quantitative phenomenon with a quantitative METT prediction.

Phenomenon 2: Typing is Slower than Reading

> Although typing is faster than reaction time, it is much slower than reading...two samples of typists averaged 246 and 259 words per minute when reading, but only 60 and 55 net words per minute, respectively, when typing. (Salthouse, 1986, p. 304)

The METT does not include a model of reading, so it cannot provide a predicted number corresponding to the 250 wpm reading rated quoted. However, we can interpret this assertion to mean that "input processes are generally not responsible for limiting the

maximum rate of typing" (Salthouse, 1986, p. 305). With a three-word look-ahead in normal transcription typing (ASSUMPTION 3), the perceptual processor is never on the critical path (except for the perception of the very first word). Therefore, the limit on the maximum rate of typing is not the perceptual process.

A sanity check is provided by the METT prediction of a theoretical maximum typing speed of 180 gwpm. This follows from the motor operator duration/skill interaction assumption, letting the motor operator time go to zero which provides an absolute upper bound. This maximum range is substantially less than the reading speed attained by many people (Salthouse's typists read at ~250 wpm on average), and, although there have been a few reported cases of world champion typists exceeding 180 gwpm, only 1% of typists ever go faster than 100 wpm!

This is a quantitative phenomenon with a qualitative METT prediction.

Phenomenon 3: Typing Skill and Comprehension are Independent

> Across typists, there is no relation between typing skill and degree of comprehension of material that has been typed. (Salthouse, 1986, p. 305)

There are no operators concerned with comprehending the perceived material in the algorithm of the METT. Comprehension is an entirely different process and, as such, would not be expected to correlate with typing speed in any way.

This is a qualitative phenomenon with a qualitative METT prediction.

Phenomenon 4: Typing Rate is Independent of Word Order

> The rate of typing is nearly the same for random words as it is for meaningful text. (Salthouse, 1986, p. 305)

The METT works on the word level and has no comprehension or syntactically high-level mechanisms, so random words will be treated no differently than meaningful text. Therefore, the METT predicts that the rate of typing would not be different for random words than it is for meaningful text.

This is a qualitative phenomenon with a qualitative METT prediction.

Phenomenon 5: Typing Rate is Slower with Random Letter Order

> The rate of typing is slowed as material approaches random sequences of letters. (Salthouse, 1986, p. 306)

The METT provides the following qualitative explanation for this phenomenon. Assumption 4 addresses the manner in which material is perceived and encoded in the METT. When the to-be-typed text is composed of words familiar to the typist, the words are stored away as chunks and accessed with the Get-Spelling cognitive operator. If the task is to type random letters, then there are no known words and they cannot be chunked as such. Therefore, the typist would have to drop down to the syllable level, where known syllables (e.g. pronounceable trigrams) could be chunked and then retrieved and unpacked. If the letters are so random that known syllables were not present, then the typist would have to drop down to the letter-level. Assuming a constant three-unit look-ahead, the perceptual processor will occasionally be on the critical path when very fast syllables are typed (e.g. alternate hand trigrams). At the letter-level, the perceptual

processor will almost always be on the critical path. Thus, typing slows down.

A more detailed analysis of this phenomenon is based on the work of West and Sabban (1982). They used progressively degraded copy to test "the effects on performance of nonlanguage as compared to language materials" (p. 372). Examples of the five different test materials are shown in Figure 5. However, only the easy prose (EP), letter combinations (LC) and letter jumbles (LJ) conditions will be considered here because they address the phenomenon in question: the effects of random letters.

West and Sabban presented their data grouped by speed of the typist performing the task; the fastest group will be used as an example. Their average speed on a standard typing test was 100 wpm, corresponding to a motor operator of 90 msec.

A task analysis reveals three different strategies for typist performing this task. All of the strategies depend on a WM limitation of three chunks (ASSUMPTION 3). The first uses a 3-word look-ahead if words are available; if not, then a 3-syllable look-ahead if syllables are available, else a 3-letter look-ahead is used (hereafter this strategy will be called *3-w-s-l*). The second possible strategy is a 3-letter look-ahead only, where the typist decided to give up trying to pronounce the letter combinations, whether they are pronounceable syllables or not (hereafter, *3-l.*). The last strategy is a 3-syllable look-ahead, if the whole word is made up of pronounceable syllables; if not, then the typist reverts to a 3-letter look-ahead (hereafter, *3-p-l*). The difference between 3-w-s-l and 3-p-l can be seen in the treatment of the word "rtleet". The former strategy would do "r", "t", "leet", and the latter strategy would do each letter separately because the whole "word" is not pronounceable. The EP material is composed of real words, so the only strategy needed is 3-w-s-l. The LC and LJ conditions might invoke the use of all three strategies. The effective typing speeds obtained with each of these strategies is shown in Table 2.

West and Sabban report their results as the percentage change between two conditions (shown in the OBSERVED column of Table 3). In order to compare results, a weighting of these different strategies must be chosen. To make a true zero-parameter prediction, the weighting would have to be chosen from considerations other than the data. In the absence of external evidence for weightings, the minimum weighting assumption (equal weights across strategies) is used. The results of these predictions are shown in Table 3.

The equal-weighting assumption leads to predictions that are quite far from the observed differences between easy prose and the progressively more random letter orderings (average absolute percent error of 43.8%). If a slightly different weighting assumption is used, namely, that for the non-words, the typist uses only 3-p-l or 3-l, equally weighted, then the fit improves (average absolute percent error of 18.1%). Thus, the METT can predict the changes in typing speeds between the West & Sabban experimental conditions, but not without additional information about the mix of algorithms actually used by the participants.

This is a quantitative phenomenon with a qualitative METT prediction.

<u>Phenomenon 6: Typing Rate is Slower with Restricted Preview</u>

> The rate of typing is severely impaired by restricted preview of the to-be-typed material. (Salthouse, 1986, p. 306)

This phenomenon has been reported by many researchers. (Hershman & Hillix, 1965; Salthouse, 1984a, 1984b, 1985; Salthouse & Saults, 1987; Shaffer, 1973; Shaffer & French, 1971; Shaffer & Hardwick, 1970). A typical task description is found in

Salthouse (1984a).

> Task 3 was to type material displayed in a single line of the video monitor and arranged such that each keystroke caused the display to move one space to the left. No visible copy was produced in this task. In successive conditions, the display contained 19, 11, 9 7 5 3 or 1 character of a 60- to 83-character sentence. The sentences were movie descriptions taken from TV GUIDE magazine and were randomly assigned to preview conditions. (Salthouse, 1984a, p. 350)

The WM limitation and perceptual chunks assumptions are particularly relevant for this task. As the preview is restricted, the three-word look-ahead is cut back to two-, then one-word look-ahead, then down to the syllable level and finally to the letter level. With a one character preview, the task is reduced to a series of choice-reaction time tasks with no look ahead.

If a whole word does not fit in the preview window, then the perceptual operator will not start until the word completed (or is shown to not to fit in the window). For instance, if the window is nine characters long, then the first view of the example sentence would be **One_reaso**. In this case, "One" would be perceived and encoded, but "reaso" would not be. The person would wait until two characters were typed and the view became **e_reason_**; "reason" would now be a whole word and a perceptual operator could work on it. Note that the perceptual operator would not start until the punctuation after the word is visible (in this case, a space) because the "reason" in **ne_reason**, for example, could easily be part of "reasonable" or other longer word.

If a word is too long to fit into the window completely, then the syllables in the window are perceived and encoded. The remaining syllables in the word are perceived and encoded as they appear until the word is complete. Separating the words into syllables for this analysis can be done from two different viewpoints: the subject's viewpoint and the viewpoint of the text. For example, with a five character window, the phrase "the elevator" would not fit; after typing the "the", the view would be **_elev**. The typist would not know what word was coming and might choose to encode this as either "e" and "lev", "el" and "ev", or "el" and "e" waiting for more letters before encoding the "v". On the other hand, the text is actually "elevator" and this is commonly broken up as "el", "e", "va", "tor" (Webster's New Collegiate Dictionary, 1979). METT analyses will always be made from the viewpoint of the text, making them simple, straightforward and objective. Thus, **_elev** would be encoded as two syllables: "el" and "e". With a three-syllable look-ahead, this sequence would not fill up working memory. Therefore, as soon as the space is typed, the view would be **eleva** and the perceptual processor would perceive and encode the next syllable, "va". After the "l" is initiated, there would be room in working memory, but the view would be **levat** and no new syllable would be visible. After the next "e" is typed, however, **vator** becomes visible and the "tor" is perceived and encoded. Then, after the "v", the view becomes **ator.** and the "." is perceived and encoded separately. Figure 6 shows the schedule chart depicting these dependencies.

An analysis was done using the example sentence, all of Salthouse's conditions and a 120 gwpm typist (Salthouse reports the data for his fastest participant, a 117 gwpm typist). The results of the analysis appear in Figure 7. The average absolute percent error is 15.8%.

This is a parametric phenomenon with a parametric METT prediction.

<u>Phenomenon 7: Alternate-Hand Keystrokes are Faster than Same-Hand Keystrokes</u>

> Successive keystrokes from fingers on alternate hands are faster than successive keystrokes from fingers

> on the same hand. (Salthouse, 1986, p. 306)

The serial/parallel processing assumption with the same-hand constraint (Assumptions 2 and 2b) imply this result. Because a keystroke cannot be initiated by the cognitive processor for the same hand until after the current keystroke is completed by the motor processor, same-hand keystrokes have a cognitive operator on their critical path that does not appear in the alternate-hand keystrokes. (See the difference between the o-n sequence and the n-e sequence in Figure 2.) Thus, the METT predicts that the interval between same-hand keystrokes will be 50 msec longer than the interval between alternate-hand keystrokes. Salthouse reports a range of 30 msec to 60 msec; 50 msec is 11.1% above the mean of this range.

This is a quantitative phenomenon with a quantitative METT prediction.

<u>Phenomenon 8: More Frequent Letter Pairs are Typed More Quickly</u>

> Letter pairs that occur more frequently in normal language are typed faster than less frequent pairs. (Salthouse, 1986, p. 306)

The METT does not try to explain this phenomenon, because the model simplifies the situation by differentiating only between hands, not between fingers. However, after the effect of same-hand vs. alternate-hand sequences is eliminated, only 4% of the variability in interkeystroke interval is accounted for by the log of the frequency of occurrence, so this is a relatively small effect when compared to the many larger effects in this list.(Salthouse, 1984b).

This is a quantitative phenomenon, although a small one, that is not covered by the METT.

<u>Phenomenon 9: Interkey Intervals are Independent of Word-Length</u>

> There is no systematic effect of word length on either the interkey interval between the space and the first letter in the word, or on the interkey interval between letters within the word. (Salthouse, 1986, pp. 307-308)

The METT works at the word level and the operator that retrieves the spelling of a word takes the same amount of time no matter what the length of the word. Therefore, the METT predicts no word-length effects.

This is a qualitative phenomenon with a qualitative METT prediction.
There is a word-length effect in discontinuous typing tasks (where a word is displayed only after the typing of the previous word is complete) (Sternberg, Knoll, & Wright, 1978; Larochelle, 1983). For the METT, this implies that the word-length effect is in the perceptual processor, not in the cognitive processor. Several researchers have found that words with more letters take longer to read (Henderson, 1982) and this may be the source of the word length effect. These reading results could easily be incorporated into an MHP-based model of discontinuous typing, but it is not necessary to introduce this complexity into the model of transcription typing because the perceptual processing is not on the critical path and thus the effect is not evident.

<u>Phenomenon 10: The First Keystroke in a Word is Slower than Subsequent Keystrokes</u>

> The first keystroke in a word in normal continuous typing is generally slower than subsequent keystrokes in the word. (Salthouse, 1986, p. 308)

The METT predicts that the first character of a word would have a longer interkey interval than subsequent letters whenever the cognitive operator that gets the spelling of the word is on the critical path. This situation occurs for all typists with speeds greater than 90 wpm. However, the average increase is very small because only one same-hand and alternate-hand character sequence brings the get-spelling operator to the critical path (i.e., the last letter of the previous word is typed by the right hand, and the first letter of the word in question is typed by the left hand). Thus, the increase in interkey interval for the first letter of a word ranges from 0% for a 60 wpm typist to 8.4% for a 120 wpm typist, for an average of 2.3%. Salthouse (1984a, 1984b) reports a 20% increase overall with a 60 wpm average typist. Thus, the METT predicts a word-initiation effect, but an order of magnitude less than the observed performance.

This is a quantitative phenomenon with a qualitative METT prediction.

The METT makes an interesting second order prediction, that the word-initiation effect increases at the high end of the expert range. This is counter to the results reported by Salthouse (1984a). He obtained about a -0.4 correlation between skill and word-initiation effect. However, his typists spanned a much larger range of skill than the METT addresses. It is possible that a mechanism different from accessing the spelling of the word exists in novice typing, causing a large word-initiation effect at the lower speed levels (e.g., perhaps a novice perceives a word, then types it, then goes back to the copy to perceive the next word). It is possible that the word initiation effect all but disappears in the low expert range, when the typist begins to look ahead of the word currently being typed, and then reappears slightly at the highest skill levels because of the word-access mechanism in the METT. Such a pattern could produce the correlation found by Salthouse. The data for different skill levels is not reported in the literature, so this pattern cannot be verified at this time. However, it is an empirically testable prediction, on the agenda for further investigation.

Phenomenon 11: The Time for a Keystroke is Dependent on the Specific Context

> The time needed to produce a keystroke is dependent on the specific context in which the character appears. (Salthouse, 1986, p. 308)

The "specific context" quoted above means the specific characters surrounding the character being typed, both ahead of and behind that character. Previously explained phenomena partially account for this effect. Phenomenon 7 (the same-hand/alternate-hand difference) phenomenon 8 (the digram frequency effect) and phenomenon 10 (the word initiation effect) all contribute substantially to this phenomenon. The METT provides explanations for phenomenon 7 and 10. However, after these effects have been controlled in careful experimentation, a context effect remains and is active over as many as three characters of context. Models that include the detailed behavior of each finger (e.g., Rumelhart & Norman, 1982) can predict these effects, but the coarse grain of the METT make this effect beyond the scope of the model. However, the locus of the explanation is probably within the motor processor and extension of the model with regards to more detailed motor programs could provide an explanation.

This is a quantitative phenomenon that is not covered by the METT.

Extending the METT to differentiate between fingers. An obvious weakness of the METT is in its inability to differentiate between fingers on a hand as we have seen in phenomena 8 and 11 (and will see in phenomenon 23). The culprit is clearly the level of approximation chosen for the METT. Given such a clear cause for the weakness, we can focus attention on improving it. One method of extending the METT with more detail, consistent with the

operating principles of the MHP, i.e., Fitts's Law, will be explored here in a re-examination of phenomenon 11.

Consider the following extension of the METT to explain data presented by Rumelhart and Norman (1982). These researchers used a detailed computer model to simulate the interkey interval between the 66 most common digraphs in a large corpus of typing data. Overall, their simulation accounts for 74% of the variance. They present actual performance data for only nine digraphs, those involving the index and middle fingers of the left hand, always ending with the striking of the *e* key. Their simulation produces estimates of interkeystroke times for these nine digraphs in "arbitrary model units" (Figure 8, p. 23), which, when I performed a regression against the observed times, account for 69% of the variance ($p < 0.01$).

According to the principles of operation of the MHP, movement time follows Fitts's Law, so, if an extension to the METT can explain this fine structure of interkey movement, it will do so with a Fitts's Law analysis. To do this analysis, a keyboard with a 0.5 in. square key, and a 0.75 in. key separation, center-to-center, was assumed (Figure 8). For same-finger digraphs, it was assumed that the finger had to move the whole way, center-to-center. For different-finger digraphs, it was assumed that the middle and index fingers stayed in the same relative horizontal position as the index finger moves to strike a key, and that striking a key on the upper or lower row moves the middle finger only half way to that row, along the trajectory it would take if that relationship were maintained. Following this assumption, the positions of the middle finger when the index finger hits another key are marked on Figure 8 with the lower-case letter of the key being hit. The straight-line distance that the middle finger had to travel was called the "Distance Moved".

It was assumed that for the same-finger transitions, the movement was initiated with a cognitive operator and movement itself could be approximated with three component moves: the finger was moved horizontally over the *e* key, then pressed down, then lifted up to be ready for the next keystroke. At this level of detail, we assumed that the up-movement overlaps with the cognitive initiation of the next keystroke. Since the *e-e* transition has no horizontal movement time, the up and down movements can be defined by the observed time to make the *e-e* transition. Our own observations of videotapes of single-finger typing reveal that down and up times are approximately symmetric and of the order of 60-100 msec. This information provides the additional constraint needed to solve for durations.

Down-time + Up-time = 165 msec   Down-time = Up-time = 83 msec

These ups and downs were subtracted from the same-finger times to get their horizontal movement time.

To get the horizontal movement time for the different-finger transitions, it was assumed that pressing the *f* key leaves the middle finger on the *d* key so the *f-e* transition involves the same horizontal movement time as the *d-e* transition. From that assumption, the down time (83 msec) and the horizontal movement time for the *d-e* transition (35 msec) was subtracted from the 168 msec observed for the *f-e* transition, to get an estimate of how much time is spent picking the left index finger up before starting to move the middle finger: 50 msec. This value is less than a pure up time (83 msec), indicating, as Rumelhart and Norman observe, movements of different fingers overlap. This estimate plus the down time (=133 msec) was subtracted from all the observed times for different finger transitions to get their horizontal movement times.

Predictions for the horizontal movement times were made by using Fitts's Law:

$$T_{pos} = I_M \log_2(2D/S),$$

where I_M is a free parameter established by regression. These predictions were regressed against the estimated horizontal movement times, forcing the regression through zero, to get an estimate of I_M, 20.70 msec.

The total predicted time is then calculated by

$T_{pred} = 20.70 * log_2(2D/S)$ + Up-time + Down-time

where D is the Distance Moved, S is the size of the key, 0.5 inches, Up-time is 83 msec for same-finger movements and 50 msec for different-finger movements, and Down-time is 83 msec.

The Rumelhart and Norman total time prediction was gotten from regressing the simulation's results against the observed times and finding the coefficient of the slope (9.17449) and intercept (104.52045) of the best-fit line and then using the equation

T(pred) = *104.52045 + 9.17449 * simulation-result.*

The observed times and the two different predicted times can be found in Table 4and Figure 9. The percent error for each key transition was calculated for both the predictions and the average absolute percent error for each prediction technique also appear. The MHP predictions based on Fitts's Law, and containing only one free parameter I_M, are slightly better than the Rumelhart and Norman predictions, containing two free parameters (slope and intercept).

When regressed against the observed interkeystroke times, these METT estimates account for 99% of the variance ($p < 0.01$). Using the data to estimate the slope of the Fitts's Law equation, produces predictions with an average absolute percent error of 2.4%.

This example of a possible extension to the METT explains the single-character context effect quite well. It does not make zero-parameter predictions, in that the slope of the Fitts's law curve was determined by regression against the data, so the predictions are qualitative, not quantitative. If the METT were to be extended to this level of detail, the next step would be to obtain performance data for other digrams and make true zero-parameter predictions.

Phenomenon 12: A Concurrent Task does not Affect Typing

> With highly skilled typists, a concurrent activity can often be performed with little or no effect on speed or accuracy of typing. (Salthouse, 1986, p. 309)

Salthouse and Saults (1987) gave typists a simple reaction-time task and a task where they had to type and perform the reaction-time task concurrently.

> The second task was an auditory reaction-time task in which responses were made by pressing a foot pedal containing a micro-switch.

> ...Task 3 was [where]...subjects typed from printed copy while responding to auditory signals with foot pedal responses. The typing task was stressed both by instructions and by delaying the introduction of the concurrent reaction-time task until subjects had typed for about 30 s. (p. 189)

The results of these tasks are presented in Table 5. The typing task was not slowed by the concurrent reaction-time task, but the reaction-time task slowed down considerably, indicating that the reaction time task had not yet become automatic and that the devices used to emphasize the typing task were successful.

The first step in an analysis of this concurrent task situation is to estimate a new motor

operator parameter for pressing a foot pedal. An algorithm for the simple reaction-time task, fit to the observed performance, gives an estimate of the new motor operator duration.

```
BEGIN
?Tone ← Get-Stimulus("Tone")                      100 msec
IF-SUCCEEDED Right-Tone?(?Tone)                   50 msec
  THEN BEGIN
        Initiate("foot-press")                    50 msec
        Execute("foot-press")                     new motor operator
        END
END
```

Subtracting the perceptual and cognitive operator times from the observed reaction-time, 269 msec, leaves an estimate of the motor operator to press a foot pedal of about 70 msec. This value happens to be the typical motor processor cycle time given by Card et al. (1983).

With this new motor operator estimate, the algorithm for the simple reaction-time task is superimposed on top of the METT algorithm for a 60 gwpm typist typing the example sentence (Salthouse and Saults' typists averaged 63.1 nwpm). The tone was assumed to start at 25 random locations within the example sentence and each location was analyzed in isolation, to simulate the random positioning of the tone within a longer portion of text (Salthouse and Saults used up to 250 words). Since the reaction time task was not automatic and the typing task was emphasize, the operators that perform the reaction time task were woven in between those of the typing task, with the typing operators taking precedence. If the perceptual processor was not busy perceiving a word when the tone started, then the perception of the tone began at the onset of the tone, otherwise the perception of the tone began as soon as the perceptual process completed the perception of the word. When the perception of the tone was complete, if the cognitive operator was not busy doing something for the typing task, the verification of the tone began, otherwise the verification waited until the typing cognitive operation was complete. Then the cognitive operators for the typing task and the reaction-time task were woven together, alternating between tasks if they were competing for cognitive processing time.[6] The motor operator to press the foot pedal began after the foot-press was initiated by the cognitive processor and the motor processor was not busy typing a character. (See Figure 10)

The concurrent reaction time was predicted from these schedule charts by measuring the time between the (randomly defined) onset of the tone and the completion of the foot-press motor operator. The average reaction time predicted by this analysis is 435 msec, 0.9% away from the 431 msec observed by Salthouse and Saults.

The effect on the average interkey interval for Salthouse and Sault's task is reported to be small. From Table 5, the mean of the interkey interval was 181 msec for the normal typing task and 185 msec when the concurrent reaction time task was added. The METT predicted an average interkey interval of 195 msec for the normal typing task, 7.7% above the observed value. The interkeystroke interval for those keystrokes that were interrupted by the reaction-time task was predicted to be 240 msec. However, this was only for those keystrokes that were directly involved in the concurrent task, not the average for the entire

[6]Sometimes the typing task involved a same-hand key sequence and the cognitive processor was waiting for the completion of the motor processor. This left enough time for the reaction-time cognitive operators to execute without alternation between tasks.

typing task. Salthouse and Saults state that the task is a combination of Tasks 1 and 2 in the same study, which they describe in detail. If the combination of tasks is taken literally, then there were 30 tones presented within a 1200 character passage. (This ratio of tones to characters was indeed the case, Salthouse, personal communication, 1988) Thus, 30 interkeystroke intervals increased to 240 msec and 1170 keystrokes remained at an average of 195 msec. With this ratio, the overall average interkey interval for typing with the concurrent task was predicted to be 196 msec, 6.0% above the observed time. This analysis supports the claim that a concurrent task has little or no effect on the typing speed of an expert typist. An interesting prediction of the METT is that the interkeystroke intervals occurring as the foot-press is occurring will increase; this prediction is left for future empirical verification.

This is a quantitative phenomenon with a quantitative METT prediction.

Units of Typing

The next five phenomena deal with various units of typing: how many characters can be typed after the copy has been removed without warning (copy span), how many characters are typed after the typist has been told to stop (stopping span), how many characters ahead of the fingers must the typist's eyes be when typing at asymptotic speed (eye-hand span), and how many characters ahead of the fingers can a change be made in the copy and still be reflected in the performance (replacement span).

Phenomenon 13: Copy Span is 7-40 Characters

> The copying span, defined as the amount of material that can be typed accurately after a single inspection of the copy, ranges from two to eight words, or 7-40 characters. (Salthouse, 1986, p. 309)

This has been measured in many different ways, and the different methods yield vastly different results, accounting for the wide range. For instance, Rothkopf (1980) measured the copy span by asking the typists to glance at the copy, remembering as much as possible, and then type it out before glancing at the copy again. This is a very different task than normal transcription typing and it yielded the result that a typist can remember up to 40 characters at a time. Salthouse (1985) measured the copy span in a way more appropriate to transcription typing, and got an average copy span of 13.2 characters for all typists (speed range from 20 to 120 gwpm), but an average of 14.6 characters for expert typists (above 60 gwpm).

The Salthouse (1985) experimental situation was as follows.

> The procedure involved presenting material on the video monitor using the leftward-moving display with a preview window fixed at 39 characters. After a predetermined number of keystrokes, the display was erased and the typist instructed to continue typing as much material as he or she was confident appeared on the display. The material consisted of eight sentences, movie descriptions from TV GUIDE magazine, with an average length of 75 characters. Two sentences each were typed with 15, 25, 35 and 45 keystrokes prior to the disappearance of the display. The median number of characters that were typed correctly after the blanking of the display served as the measure of copying span. (p. 267)

The METT can easily model this task. Predictions can be made at several different levels of detail. First, a quick and dirty analysis is presented, then a more detailed analysis of the Salthouse task.

On average, a word is 5 letters long. If there is a 3-word look ahead, there is, on average,

15 letters in the perceptual buffer, or working memory. If the display is removed randomly, it will be removed, on average, 2.5 characters into a word. Therefore, there will be a copy span of about 2.5 words, or 12.5 characters. This prediction is within 14.4% of the observed 14.6 characters.

To do the more detailed analysis, more detailed information is required:

> Net typing speeds for the 29 typists ranged from 18 tp 113 NWPM with a mean of 62.4. Gross speeds ranged from 20 to 120 words per minute with error percentages from 0.1 to 4.2.
>
> ...four speed groupings (six typists at less than 40 NWPM, M=27.8; seven typists at between 40 and 60 NWPM, M=48.1; eight typists at between 60 and 80 NWPM, M=72.5; and eight typists with speeds greater than 80 NWPM, M=90.6).
>
> T...he correlation between NWPM and the median number of characters typed after the disappearance of the test display (i.e., the copying span) was .35 (.10>p>.05). Across all typists the copying span averaged 13.2 characters, and from the slowest to the fastest speed groups the spans averaged 10.5, 12.4, 15.5, and 13.6 characters, respectively, $F(3,25) = 1.75$, $p>.15$. (pp. 267-268)

Salthouse used different sentences and four different stopping points, but the same effect can be gotten by using only the example sentence imposing a stop after every character, and averaging them all together. Assuming an 80 wpm typist (the approximate average of the typists in the expert range, above 60 wpm), the copying span was predicted as if the display disappeared at each position between character 1 ("O" in "One") and character 75 ("r" in "for"); after character 75 there are no more characters to see with look-ahead. It is assumed that there is a three-word look-ahead (WM limitation and perceptual chunks assumptions) and as soon as the last character of a word (including the punctuation and space after it) is out of working memory, i.e., the cognitive processor has sent a signal to the motor processor with an initiate-letter operator, the next word can be perceived.

The disappearance of the copy is triggered by the typing of a character, which is the completion of a motor operator. Since the initiation of characters by the cognitive operator triggers the perception of the next word, the perception of the word takes a finite amount of time, and the cognitive and perceptual processors work in parallel with the motor processor, the relationship between the character just typed and the contents of working memory is not a straightforward one. The relationship is determined by the same- and alternate-handed history of the text being typed and the duration of the operators. A *task timeline* [7] for the sentence being typed is necessary to chart the contents of working memory at every stopping point. The copy span is how many letters are typed after the copy disappears, so it is a combination of the letters that have already been initiated by the cognitive processor (but not yet executed by the motor processor) and what is left in working memory, which can be initiated and executed.

Figure 11 shows a small portion of a task timeline from which the copy span can be determined. Table 6 shows the letters that can be typed for several different stopping points. The copy span for the average typist in the expert range (~80 gwpm), from this detailed calculation is 11.9 characters (19% from the observed 14.6 characters).

---

[7] A task timeline is an alternative representation of a schedule chart. Time is represented along the horizontal. The width of the task boxes are proportional to their duration. Tasks that overlap in time appear stacked above each other. The critical path is indicated in the task timeline by boxes that have no striped area at their right side. The striped area is the *slack time* for each subtask, the time that the subtask could be delayed without affecting the duration of the total task. Subtasks on the critical path have no slack time, and thus no striped area.

This is a quantitative phenomenon with a quantitative METT prediction.

The first, quick and dirty analysis gives a good prediction of the copy span, in fact, it is even slightly better than the prediction resulting from the more detailed analysis. Why should more effort be spent to do the more detailed analysis? For many purposes (e.g. designing a human-computer interface), there is no reason to do the more detailed analysis. For purposes of developing a model of typing, it should be demonstrated that the mechanism of the model does not get in the way of good predictions. If situations occur where the entire mechanism of a model is not necessary for prediction, then the quicker predictions should be used, but more detailed analyses should not be substantially worse. The copy span task is one situation where considering only the model of working memory in the METT suffices to produce a good prediction. However, this simple analysis is not detailed enough to predict most of the Salthouse 29. For example, it would predict that the copy span is constant over skill. The more detailed analysis procedure, however, produces different predictions with a range of skills, reproducing the pattern observed in actual data (see Phenomenon 28).

Phenomenon 14: Stopping Span is Between 1 and 2 Characters

> The stopping span, defined as the amount of material to which the typist is irrevocably committed to typing (Logan, 1982), averages only one or two keystrokes. (Salthouse, 1986, p. 309)

Logan (1982) measured this span directly by asking typists to stop typing as soon as they heard an auditory stop-signal. He did this in three slightly different experiments. The first experimental task, a time-contingent, discontinuous typing task, was as follows.

> Three-, five-, and seven-letter words were centered on the screen...The words were exposed for 1,000 msec, preceded by a fixation point that was exposed for 500 msec and was extinguished immediately before the word appeared. The intertrial interval was 2,000 msec and began as soon as the word was extinguished...The stop signal was a 500-msec, 900 Hz tone...It was presented at one of four delays (500, 650, 800, and 950 msec) following the onset of the word...subjects had no visual record of what they typed...The words within each length condition were balanced for hand repetition and alternation in the keystrokes they required...(Logan, 1982, p. 780)

To analyze this task, schedule charts were constructed for the perception and typing out of three- five- and seven-letter words with all possible same-hand, alternate-hand sequences. Imposed on top of these schedules was a perception of the tone (assumed to be 100 msec, estimated from a click-counting experiment, see Card et al., 1983, p.33) and a cognitive operator that recognized the tone to be the stop signal. The start of the perception operator was positioned at the start of the stop-signal (500, 650, 800 or 950 msec). The cognitive operator followed immediately and prevented any more characters from being initiated after that point in time. The number of characters typed after the stop-signal started was then the number of characters that had been initiated by the cognitive processor before the cognitive operator recognizing the stop-signal had begun. The motor operators associated with the initiation cognitive operators then completed the act by typing out the characters. The timeline form of the diagram shows the sequence of events most clearly (See Figure 12).

The average stopping span predicted by this analysis is 1.76 characters, 12.1% above the 1.57 characters observed by Logan. Although the METT is a course-grain model of typing and is intended to capture only first order effects, it is interesting to look more closely at the data and see which patterns the METT can reproduce. Logan reports several patterns in the data that suggest that typists do not type whole words before they stop, a behavior also not predicted by the METT.

> If they had typed whole words, the mean span should have decreased with stop-signal delay because fewer letters remain to be typed at the longer delays, and the spans should increase with word length

> because at each stop-signal delay, more letters remain to be typed with longer words. However, the data generally disconfirm these predictions. For five- and seven-letter words, the span increased with stop-signal delay instead of decreasing, and it did not increase with word length. For three-letter words, span increased from the 500-msec delay to the 650-msec delay but decreased at the longer delays when subjects had nearly completed the word before the signal occurred. This resulted in a lower mean span for three-letter words (1.43 letters) than for five- and seven-letter words. (Logan, 1982, p. 781)

Most of these patterns were predicted by the METT. For five- and seven-letter words the span did not decreased with delay, and did not differ with word length. For three-letter words, the METT did not predict the rise in span between the 500-msec and 650-msec delay, but the model did complete typing the word in many cases before the 950-msec delay, causing a substantially lower average span for that condition.

Two other patterns exist in the data:

> ...assuming that the span reflects the latency of the (internal) response to the stop-signal and that the latency of the response to the stop signal is constant over stop-signal delay...we would expect subjects to type fewer letters when the stop signal occurred at the 500- and 650-msec delays before subjects began typing...[Another pattern is that] the span appears to increase as subjects progress through the response sequence, although this effect did not replicate in [the other experiments]. (Logan, 1982, p. 782)

The first pattern did not appear in the METT predictions, span remained the same or decreased between the 500-msec and 650-msec delays. However, if Logan's explanation is correct, that people had not yet started typing before these delays occurred, the METT could not possibly have picked this up because the predicted latency between presentation of the word and being committed to typing the first letter was always one perceptual operator and two cognitive operators (to get the spelling and initiate the first letter), 440 msec, less than the 500 msec of the smallest delay. This indicates that the perceptual process may be the weakest part of the METT. Since the other pattern in Logan's data did not replicate, the fact that is not predicted with the METT speaks in the model's favor.

Logan's second experimental task made the stop-signal contingent on an event, the typing of a specific character, and was as follows.

> ...the same as in Experiment 1 except that the routine that accepted responses from the keyboard was rewritten to present a stop signal when a prescribed number of keystrokes had been registered (i.e., immediately after the *n*th keystroke). The copy to be typed was the five- and seven-letter words from the first experiment...The stop signal occurred on 20% of the trails...at one of four delays (after 1, 2, 3, or 4 keystrokes had been registered). (Logan, 1982, p. 782)

This event-contingent stopping task was analyzed in the same way as the time-contingent task, looking at all possible same- and alternate-hand sequences of five- and seven-letter words and all possible stop-signal onsets (after the first, second, third and fourth letters typed). The average predicted stopping span was 1.55 characters, 9.9% above the observed 1.41 characters. The observed results showed that

> ...the span was not substantially affected by word length or stop-signal delay except when signals occurred on the fourth letter of five-letter words. In this situation, only one letter was left to be typed. Clearly, there was no tendency to type whole words. (Logan, 1982, p. 783)

This pattern was reproduced in the predictions.

Finally, Logan's third experiment, also event-contingent, examined stopping behavior within a sentence rather than a single word. The experimental task was as follows.

> The copy to be typed was a set of 300 sentences of the form "the [noun] [verb]ed the [noun]," made from the five- and seven-letter words from Experiments 1 and 2. The nouns were either 5 or 7 letters long, but the verbs were 5, 6, 7, or 8 letters long because letters sometimes had to be added to make sense...There were 300 trials, and the stop signal occurred on 33% of them...at one of 20 positions in the sentence...Twelve of the positions were within words (the first, second, third, and fourth letters of the noun, the verb, and the second noun), and eight of the positions were between words (the last letter and the following space from the first four words of the sentence). (Logan, 1982, p. 784)

The same type of analysis was used for this task and the result was an average predicted stopping span of 2.08 characters, 3.7% below the observed average of 2.16 characters. In this experimental situation, the METT predicted fewer of the patterns. However, the pattern that Logan discussed at greatest length was predicted by the METT. The pattern is that "Spans before 'the' tended to include the word and the space following it" (p. 786). Logan cites this as evidence that "the" is typed ballistically and attributes it to the frequency of the word in the language and its frequency within the experimental situation as well. The METT suggests another explanation. The letters and the space in "the_" alternate hands, this allows the cognitive processor to initiate the whole word and the space in advance of the motor execution of the letters. Thus the stop signal often comes at a point where the letters have been initiated, and must ballistically complete. This situation also exists with other frequent words, like "and_", and is an interesting prediction for future empirical verification.

Salthouse presents this as a quantitative phenomenon. However, examination of Logan's experiments allows us to upgrade it to a parametric phenomenon, because the different tasks produce different estimates of the stopping span (see Table 7). The METT predicts the pattern of results found by manipulating the tasks with an average absolute percent error of 8.6%.

This is a parametric phenomenon with a parametric METT prediction.

<u>Phenomenon 15: Eye-Hand Span is Between 3 and 8 Characters</u>

> The eye-hand span, defined as the amount of material intervening between the character receiving the attention of the eyes and the character whose key is currently being pressed, ranges between three and seven characters for average to excellent typists. (Salthouse, 1986, p. 310)

This phenomenon has been measured in two ways. Butsch (1932) recorded eye-movements while a person was typing and synchronized them with keying records to determine the position of the eye at each keystroke. He reported that the eyes are 4.9 characters ahead of the fingers, on average, for a group of typists averaging 60 wpm. The other method of estimating the eye-hand span is to use the restricted preview paradigm (see Phenomenon 6) and define the eye-hand span as the smallest number of characters in the preview window at which the typists reaches her asymptotic speed. Using this method, several researchers (Hershman and Hillix, 1965; Salthouse, 1984a, 1984b, 1985, and Shaffer, 1973) have reported between three and eight characters as the eye-hand span for moderately skilled typists.

Returning to the detailed examination of the restricted preview task in Phenomenon 6, the figure (Figure 7) indicates that the typist was observed to reached her asymptotic speed with a window between seven and nine characters in length (Salthouse assumes this to be an eight-character eye-hand span). The predicted curve is much smoother and actually does not reach normal typing speed until the 19-character window. This discrepancy can be partially accounted for by the difficulty in comparing the predicted results to the criterion used by the researcher to determine when asymptotic speed had been obtained.

> The eye-hand span in Study 1 was defined as the smallest window at which the first quartile was greater than the second quartile of normal typing. This procedure effectively identified the span as the number of display characters at which 75% of the interkey intervals exceeded the median interval from normal typing. (Salthouse, 1984a, pp. 354-355)

A revised definition of reaching asymptotic speed with the predicted performance uses an estimate of the observed distribution. That is, the observed data indicate that the first quartile, is on the average, 19% smaller that the second quartile. Thus the predicted performance was taken as the median, or second quartile estimate, and the first quartile was estimated at 19% smaller than that value. That gives the predictions in Table 8.

These predictions of the first quartile show that the first quartile becomes equal to the second quartile of normal typing (96 msec) at a nine-character preview window. Thus, a nine-character eye-hand span is predicted using this analysis, 12.5% above the eight-character eye-hand span reported for this subject.

This is a quantitative phenomenon with a quantitative METT prediction.

Phenomenon 16: Eye-Hand Span Decreases with Decreasing Meaning

> The eye-hand span is smaller for unfamiliar or meaningless material than for normal text. (Salthouse 1986, p. 310)

The METT analysis of this phenomenon is obtained from the critical path on the schedule chart for typing random letters with different preview windows. With infinite preview (Figure 13), the critical path is dominated by the perceptual processes, and will not change at all until the window is at 2-letters. At a 2-letter window, one motor operator appears on the critical path because the window must be advanced by the typing of a character. This one slight change increases the average interkey interval by only 3% (Figure 14). When the preview window is decreased to 1-letter, then the critical path changes drastically, becoming completely serial, every operator dependent on the completion of every previous operator (Figure 15). The average interkey interval increases by about 64% over the 2-letter window interval. The quartile distribution reported by Salthouse again puts the first quartile 19% below the second quartile for this random-letter condition. Assuming this distribution, the predicted eye-hand span is somewhere between one and two letters; predict 1.5 letters given no other information. Comparing this to the reported 1.75 characters, there is a 14.3% absolute error between predicted and observed eye-hand span for typing random letters.

This is a quantitative phenomenon with a quantitative METT prediction.

Phenomenon 17: Replacement Span is about 3 Characters

> Typists appear to commit themselves to a particular character approximately three characters in advance of the current keystroke. (Salthouse, 1986, p. 311)

The replacement span is the number of characters between the character currently being typed and where a change can be made in the copy, which the typist will be able to detect and typed as changed. Salthouse and Saults (1987) measured the replacement span as follows.

> ...the replacement span is defined as the key-stroke-replacement interval corresponding to a .5 probability of typing the second [replaced] character....Salthouse and Saults (1987) found the replacement spans to average 2.8 and 3.0 characters in two studies... (Salthouse, 1986, p. 311)

The task given to the typists was:

> In this task the material will always be lower case, but on some occasions a letter will be changed in the display. You should ignore the original character when this happens and type the 'corrected' version that appeared most recently on the display. Remember to try to type exactly what appears on the screen in as normal a fashion as possible. (Salthouse & Saults, 1987, p. 189)

An analysis of the example sentence laid out for a 60 gwpm typist (approximately the average typist in the study) was done to simulate this task. The task timeline form of the diagram was used. Excluding the first three and the last 13 letters (the first three arbitrarily because I wanted the typing to get going before switching characters and the last 13 because those characters were beyond where there was anything more to look-ahead and see), each letter was examined to find out when the change had to occur in order for the typing of that letter to be stopped, and changed to the new letter. For a letter to be stopped, a cognitive operator recognizing a stop signal had to be started before the cognitive operator that initiates the typing of the letter was started. For this cognitive operator to begin, there would have had to be a perceptual operator (perceiving that a change in the display had occurred) started and completely finished. This perceptual operator did not have to perceive the exact nature of the change, just that a change had occurred. One hundred msec was used as the duration of this "perceive-a-change" perceptual operator, because it is a typical perceptual operator (Card et al.'s Middleman perceptual processor cycle time). This point in time (see Figure 16) was the point at which the change had to occur in the stimulus for the typist to make the change in her typing, with one important constraint. If the perceptual processor was busy elsewhere (looking ahead at the copy) then the change would not be noticed, so the change would actually have to occur before the perceptual processor looked ahead. This constraint is meaningful within the model; the processors work serially within themselves. It is also meaningful within a detailed task analysis; if the person is looking ahead, then she might not notice a change in a different part of the display. After finding the necessary starting point of the change in the display, a count was made of the motor operators that would complete themselves in between this starting point and the point that the changed-letter was to have been typed. Because the METT has no mechanism for figuring the probabilities of behavior, this count is used as the replacement span. With this analysis, the replacement span was predicted to be 2.1 characters, 27.6% below the observed 2.9 characters.

The analysis of the replacement span is very similar to the analysis of the stopping span. The difference between the two tasks is that in the stopping span task, the perceptual process can be completed before the stop-signal occurs, whereas in the replacement span task, perception of the to-be-typed copy occurs throughout the task. The stopping span predictions are very good, even parametric with slightly different tasks, and the replacement span prediction is not as good. This indicates that the model of perception is probably the weak point of the METT. However, qualitatively, the replacement span is predicted to be positioned between the stopping span and the eye-hand span, as it occurs in the observed data. Therefore, this is a quantitative phenomenon with a qualitative METT prediction.

Errors

The next five phenomena deal with errors in typing. In its current form, the METT makes only response-time predictions, and cannot predict the types of errors that will be made. Therefore, all these phenomena are beyond the current scope of the METT. However, the error phenomena are presented here because they provide constraints on the type of extensions that might be made to the METT. The implications of these phenomena for the METT are reviewed in the Discussion.

Phenomenon 18: Only a Fraction of Errors are Detectable Without Reference to the Typed Copy

> Only between 40% and 70% of typing errors are detected without reference to the typed copy. (Salthouse, 1986, p. 311)

Additional detail is supplied by Rabbit (1978), where he reports the types of errors detected when typists cannot see what they have typed (Table 9). In his terminology, "compound errors" refers to when "two or more letters of text might be transposed or garbled in order, or an incorrect keystroke, which had no relation to the text, might be followed by one or more other incorrect keystrokes, by a transposition of two letters, or by an omission of one or more letters" (p. 949). Unfortunately, the compound and single-letter categories cannot be mapped into the substitution, intrusion, and transposition errors reported elsewhere and discussed by Salthouse. Thus, the only clear conclusion available from the Rabbit results is that omission errors have a very low probability of detection. In addition, Rabbit states (though without presenting numeric support) the fact that the typists "made compound errors due to `de-referencing' of hands with respect to the keyboard, and the probability of detection of such errors was, not surprisingly, very low" (p. 949).

Phenomenon 19: Substitution Errors are Mostly Adjacent Keys

> Many substitution errors involve adjacent keys....Results from highly skilled typists indicated that from 31% to 59% of substitution errors involved horizontally adjacent keys, and between 8% and 16% involved vertically adjacent keys. (Salthouse, 1986, p. 312)

Phenomenon 20: Intrusion Errors are Mostly Short Interkey Interval

> Many intrusion errors involve extremely short interkey intervals in the immediate vicinity of the error....[This is] interpreted as being caused by the nearly simultaneous contact of two adjacent keys by a finger imprecisely positioned above the target key....the median ratios [are] considerably less than 1.0 for the error keystrokes [0.68]...and the immediately following keystroke [0.87]... (Salthouse, 1986, p. 313)

Phenomenon 21: Omission Errors are Mostly Long Interkey Interval

> Many omission errors are followed by a keystroke with an interval approximately twice the overall median....[This is] consistent with insufficient depression of the keystroke for the omitted character such that its latency is incorporated into the interval for the following keystroke. (Salthouse, 1986, p. 313)

Phenomenon 22: Transposition Errors are Mostly Cross-Hand

> Most transposition errors are cross-hand rather than within-hand....The percentage of total transposition errors that involved fingers on opposite hands reported by Grudin was 78%, compared with a chance value...of approximately 53%. (Salthouse, 1986, pp. 313-314)

Skill Effects

The last seven phenomenon deal with effects of skill in typing. The METT is a model only of expert transcription typing, so the skill effects are examined only in the range above 60 gwpm. Some of the skill effects are still present in this range, but less pronounced than with the full range of novice through expert typists.

Phenomenon 23: 2-Finger Digrams Improve Faster than 1-Finger Digrams

> Digrams typed with two hands or with two different fingers of the same hand exhibit greater changes with skill than do digrams typed with one finger. (Salthouse, 1986, p. 314)

The reported decreases in interkeystroke interval vary. Salthouse (1984a) reports rates of -2.08 msec/nwpm for two-hand digrams, -2.38 msec/nwpm for two-finger digrams, -1.91 msec/nwpm for one-finger digrams, -0.85 msec/nwpm for one-letter digrams. The METT does not cover this phenomenon because it does not differentiate between fingers. This is a quantitative phenomenon not covered by the METT.

Phenomenon 24: Tapping Rate Increases with Skill

> The rate of repetitive tapping is greater among more skilled typists. (Salthouse, 1986, p. 314)

Increase in typing skill decreases the motor operator for hitting a key (motor operator duration/skill interaction assumption). It is assumed that the motor operator for the similar task of hitting a single key in repetitive tapping would also decrease (operator similarity assumption). Thus, the speed of the tapping would increase with typing skill. However, no rates of change of tapping interval are available given in the literature to test quantitative predictions.

This is a qualitative phenomenon with a qualitative METT prediction.

Phenomenon 25: Variability decreases with skill

> The variability of interkey intervals decreases with increased skill of the typist. At least two types of variability can be distinguished in typing, and both have been reported to be smaller among fast typists. One type is interkeystroke *variability, in that it refers to the distribution of interkey intervals* across different keystrokes and different contexts. The second type of variability is intrakeystroke variability or repetition variability. This is the distribution of interkey intervals for the same keystroke in the same context, but across multiple repetitions. (Salthouse, 1986, p. 315)

In as much as the METT does not provide the detail needed to predict the second variability, it also does not make any predictions concerning intrakeystroke variability changing with skill. However, with interkeystroke variability, a strong prediction comes from the same-hand constraint assumption (Assumption 2b) and the motor operator duration and skill interaction (Assumption 6c). The same-hand constraint assumption implies that the difference between a same-hand keystroke and an alternate-hand keystroke will be the time of a single cognitive operator, 50 msec. The motor operator duration and skill interaction assumption implies that this will be a constant difference across the range of skill. Thus, the METT makes the prediction that the interkeystroke variability, when considering same-hand and alternate-hand keystrokes, will not decrease with increasing skill.

This prediction seems to be contrary to the phenomenon as stated by Salthouse. However, closer examination of the data indicates that the decrease in variability is primarily due to same-finger digrams (Gentner, 1983; Salthouse, 1984a, 1986). The METT does not differentiate between fingers on a single hand, so it does not make a prediction about the variability within a hand or how it changes with skill. The prediction that the absolute difference between same-hand and alternate-hand keystrokes will be constant across skill and about 50 msec finds strong support in the data reported by Ostry (1983) (Figure 17). He shows a constant difference between same-hand and alternate-hand keystrokes of about 45 msec (giving an error of 11.1%). Given the good agreement between data and prediction for the type of variability the METT can address, this will be considered a

successful quantitative prediction.

This is a quantitative phenomenon with a quantitative prediction.

Phenomenon 26: Eye-Hand Span Increases with Skill

> The eye-hand span is larger with increased skill. (Salthouse, 1986, p. 315)

The eye-hand span is the minimum number of characters ahead of the character being typed that the eye must be to achieve normal typing speed. Using the analysis technique described in Phenomenon 15, the eye-hand span was calculated for a 120 wpm typist examined in that section and for 60 wpm and 90 wpm typist (Table 10). Assuming the same distribution as the observed data, with the first quartile being 19% smaller than the second quartile, then the three-character preview window condition is the first condition where the interkeystroke interval of the first quartile exceeds the second quartile interkeystroke interval of the infinite window condition. Therefore, the eye-hand span would be set at 4 characters for 60 gwpm typists, at 8 characters for a 90 gwpm typist, and at 9 characters for a 120 gwpm typist. The best fit line to these results has a slope of 1 character per 12 wpm increase in speed (0.083 characters per gwpm speed-up). Salthouse reports slopes between 0.025 and 0.060 characters per wpm speed-up, the predicted result being 38.3% above the high end of this range.

This is a quantitative phenomenon with a qualitative METT prediction.

Phenomenon 27: Replacement Span Increases with Skill

> The replacement span, indicating how far in advance of the current keystroke the typist commits to a particular character, is larger among more skilled typists. *(Salthouse, 1986, p. 315)*

Using the analysis technique described in Phenomenon 17, the replacement span was predicted for three typing speeds within the expert range (60 gwpm, 90 gwpm, and 120 gwpm). The replacement spans were 2.1 characters for the 60 gwpm typists, 2.5 characters for the 90 gwpm typists, and 3.2 characters for the 120 gwpm typists, showing an increase of about 1 character for every 60 gwpm increase in speed. This slope is half the 1 character/30 wpm speed increase reported by Salthouse. The original paper reporting this results (Salthouse & Saults, 1987) does not show a scatter plot of the observed performance, so it is not possible to tell whether the slope flattens out at the expert range of the skill dimension. This is another prediction of the METT left for future empirical verification.

Given the difference between predicted and observed slopes, this is a quantitative phenomenon with a qualitative METT prediction.

Phenomenon 28: Copy Span is Dependent on Skill

> The copying span is moderately related to typing skill. (Salthouse, 1986, p. 315)

The copy span is the maximum number of characters beyond the character being typed that a typist looks in the course of normal typing. Using the analysis technique described in Phenomenon 13, the copy spans for the average typists in each of the two expert speed ranges (between 60 and 80 nwpm, M=72.5 nwpm, and over 80 nwpm, M=90.6 nwpm) were calculated. For 70 gwpm typists, the copy span is 12.2 characters, 21.3% from the observed 15.5 characters. For 90 wpm typists, the copy span was 11.7 characters, 14.0% from the observed 13.6 characters, for an average error of 17.7% for the two predictions.

Copy span is observed, and predicted, to be dependent on skill within the expert range. However, the copy span goes down with skill rather than up, contrary to the general trend across the whole range of novice to expert typists. The specific mechanism for this trend within the METT is a complex relationship between the durations of different operators, observable through analysis of detailed timeline diagrams. The intuitive statement of the meaning of this downward trend in the expert range is that as skill increases, not as much need be kept in working memory to maintain high levels of speed.

This is a parametric phenomenon with a parametric METT prediction.

Phenomenon 29: Stopping Span Increases with Skill

> Fast typists have larger stopping spans than slow typists. (Salthouse, 1986, p. 315)

Logan (1983) reports no significant relationship between stopping span and skill. However, in a later experiment with more subjects, Salthouse & Saults (1987) report a positive correlation between stopping span and typing skill. The METT prediction of this relationship was examined by calculating the stopping spans for typists at 60 gwpm, 90 gwpm, and 120 gwpm for Logan's event-contingent experimental task (see Phenomenon 14). The stopping spans were 1.55 msec, 1.55 msec and 1.76 msec, respectively. Thus, the METT does not predict a strong increase in stopping span with skill, although a slight net increase is indicated. Salthouse and Saults do not report the slope of the relationship (just the correlation coefficient), making this a qualitative phenomenon.

Thus, this is a qualitative phenomenon with a qualitative METT prediction.
...and one more for good measure: Detection span

After the publication of the Salthouse 29, Salthouse and Saults (1987) defined a new span characterizing typing behavior, the detection span. The detection span is "defined as the median number of characters intervening between the target and the character currently being typed"
(p. 190).

The instructions for the task measuring the detection span were as follows.

> In this task a large number of characters will always be visible on the display, but occasionally a capital letter will appear. Whenever you notice a capital letter anywhere on the line you should press the '/' key as soon as you can and then resume typing. The capital letters should not be typed as capitals, but whenever you detect an upper-case letter you should press the '/' key. Always try to type as normally as possible. (Salthouse & Saults, 1987, p. 189)

Using this task, with a stimulus material of randomly arranged four-letter words, the detection span was defined as the median number of characters between the target and the current character, however, the mean of the detection span was also reported and that is the measure the METT predicts. The observed mean detection span was 8.1 characters (SD = 4.6 characters) for one study and 7.8 characters (SD = 5.0 characters) for another. Both sets of participants had an average typing speed of about 60 gwpm.

The distribution of the observed detection spans were quite flat. Salthouse and Saults offer this interpretation of that result:

> The magnitude of the detection span was quite variable across subjects, possibly because several different strategies could be used in this task. On the one hand, subjects could simply try to type normally and emit a detection response only when a target was accidentally encountered near the

> occurrence of one's current keystroke. On the other hand, the subject could periodically decide to interrupt his or her typing to scan for targets, thus detecting the target at very great distances and resulting in larger detection spans. (Salthouse & Saults, 1987, p. 193)

This multiple strategy approach is inherent in the METT, and three different algorithms were assumed in order to make the detection span prediction.

The first algorithm (called the *spelling algorithm*) corresponds to Salthouse and Saults' first strategy. The algorithm assumes that the perception operator includes an encoding of whether a capital exists in the spelling of the word. When the spelling of the word is brought into working memory in preparation for that word being typed, a test is performed on the spelling to see if a capital exists. If a capital does exist, then the result of the test is to type the '/'. A critical path analysis of this algorithm (Figure 18) reveals that, since the initiation of the preceding space triggers the retrieval of the spelling of the next word, the space is always typed before the '/' is hit. Since the letters of the word containing a capital cannot be typed until the '/' is hit, then the position of the capital letter in the word is the detection span (i.e., if the capital is the first letter, the detection span is one; if it is the second, the detection span is two, etc.). With a stimulus material of only four-letter words, this gives detection spans of 1, 2, 3, and 4 characters.

The second algorithm (the *perception algorithm*) has two versions: (*wait* and *parallel*. The assumption is made that the participants obeyed the instructions to "Always try to type as normally as possible" (p. 189), and did not stop and look ahead of where they would look in the course of normal typing, but made the judgment about whether a capital existed after the word was perceived normally. It is assumed that the perceptual operator encoded whether a word included a capital independent of its spelling. The test for a capital was performed as soon as the perception was complete, and the '/' hit if a capital was detected. The difference between the two versions of the algorithm is that in the wait version, a perception of a word is initiated in the course of normal typing, but the cognitive processor waits for the perception to be complete and makes the test before continuing with typing. Thus, this wait makes the perception-wait algorithm display behavior similar to the stop-and-scan strategy that Salthouse and Saults proposed. In the parallel version, the perception proceeds as it does in normal typing and the cognitive processor continues with the typing task until the perception is complete, and then the test is made.

In the wait case, a critical path analysis (Figure 19) reveals that since the initiation of a space at the end of a word is the cognitive activity that triggers the perception of the next word, and the decision about the capital waits for the perception to be completed, then the first character of the word that is two words before the word with the capital will always be the character at the leftmost side of the screen and the detection span will be 10 plus the position of the capital within the word (i.e., 11, 12, 13, or 14).

In the parallel case, the critical path analysis (Figure 20) is more complex because the same-hand/alternate-hand pattern of letters influences how many characters are typed during the perceptual process. The detection spans predicted using this algorithm range from 7 to 13 characters.

Averaging over these possible algorithms, yields a predicted detection span of 8.56, 7.7% away from the 7.95 average observed by Salthouse and Saults. It is also worth noting that the extreme detection spans predicted are 1 and 14, whereas the extremes observed are 1 and 15. This result suggests that a true stop-and-scan strategy, like the one proposed by Salthouse and Saults, that assumes a violation of the task instructions by the participants, is not necessary to explain the data.

This is a quantitative phenomena with a quantitative prediction, but since it was not a part of the original Salthouse 29, it will not be continued through the assessment of the METT's performance on its run of that gauntlet.

SUMMARY

This article explored the predictive power of the METT, a model of transcription typing based on the MHP framework. The basic form of the model and two parameters (time to perceive a written word, syllable or letter and the cognitive cycle time) were transported from research associated with an MHP model of S-R compatibility to the typing domain. A few task-specific assumptions were added to make the MHP applicable to typing tasks. The motor operator parameter for expert typists of different average typing speed was calculated from performance on a standard typing test. Several other time parameters (e.g., perceiving an auditory tone) were estimated from previous experiments in other domains (e.g., the fusion of clicks experiment, Cheatham and White, 1954). From these a priori definitions and estimates, the METT was used to analyze the Salthouse 29 phenomena of typing.

The Salthouse 29 phenomena range from being qualitative in nature (e.g., Phenomenon 3: There is no relation between typing skill and degree of comprehension of material that has been typed.), to being quantitative and even parametric (e.g., see Figure 7 in Phenomenon 6, restricted preview). Likewise, the METT makes predictions that agree parametrically, quantitatively, or qualitatively with the reported phenomena. Table 11 gives the totals of the different types of accounting the METT makes of the data. A summary of the way in which the METT accounts for the Salthouse 29, by individual phenomenon, is given in Figure 21.

DISCUSSION

This presentation of the METT is an example of Newell's the vision of appropriate psychological research. Newell suggested three paths to guide psychological research away from local, empirical investigation of phenomena and binary questions, toward a more mature science with theories that provide explanations, predictions, and prescriptions for control. The first path was to study a complex task in full. This path was taken, collectively, by dozens of typing researchers and Timothy Salthouse in his 1986 review. The Salthouse 29 are a clear example of what Newell had in mind for describing the robust phenomena of a complex task. They delineate the outline and much of the substance, for typing, of what Newell has called *The Great Data Puzzle* (Newell, 1990) and allow additional pieces of data to be fit into the picture, or be judged spurious. As Salthouse says, "...one can at least be assured that the phenomena being explained by various theoretical models are genuine, and do in fact require explanation" (p. 304). The other two paths, constructing complete processing models and applying a model of a single phenomenon to lots of diverse situations where that phenomenon arises, have been followed by the METT and subsequent work stemming from the METT. The benefits of following these two paths will be discussed with respect to the performance of the METT itself and its contributions to creating a unified theory of cognition and its contributions to the design of computer systems.

Constructing complete processing models

Newell's second path for research was to construct complete processing models. For typing, completeness means modeling the perception, cognition, and motor behavior involved in performing a typing task, in enough detail to constrain the model's prediction of the methods employed in the performance of a task. The METT is clearly on this path.

Starting with the MHP framework, and adding the six specific assumptions derived from a task analysis of transcription typing, the METT defines a typing machine: insert a specific task, turn the crank, and out comes a performance prediction. This paper has been an exhaustive exercise in turning the METT's crank for the Salthouse 29.

One benefit of such a machine is that it becomes independent of the theorist, and is open to objective examination. The six assumptions of the METT leave very little room for interpretation of the method used to accomplish each specific typing task. Often only one method emerges (e.g., phenomena 1,4,7,9, etc.), sometimes a few (e.g., phenomenon 5). The predictions are based more on the mechanics of the model than on the skill of the theorist. This allows objective inspection of the model for its strengths and weakness, and how it can be modified and improved.

Examining Models for Strengths and Weaknesses

One strength of the METT lies in its ability to predict performance time. Most of the phenomena that involve time to perform a task, or spans derived from performance time, are extremely well predicted. This strength comes from the well-defined information-flow control structure and reasonable estimates of operator duration.

Another strength is in the METT's ability to predict irrelevant factors in typing, like comprehension (phenomenon 3), word-order (phenomenon 4), and word-length (phenomenon 9). This strength arises directly from the examinability of the model, e.g. there is no comprehension mechanism in the METT, so the METT predicts that comprehension is irrelevant to typing skill.

A weakness of the METT is in its inability to differentiate between fingers on a hand (phenomena 8, 11, & 23). The culprit is clearly the level of approximation chosen for the METT. Given such a clear cause for the weakness, we can focus attention on improving it (and we began to do just that in the discussion of an extension to the METT based on Fitts's Law in Phenomenon 11).

The most obvious weakness is that the METT cannot predict errors. When you put typing tasks into the METT and turn the crank, only perfect, error-free behavior comes out. Again, the mechanistic form of the METT allows us to examine the sources of this weakness and focus future attention on the mechanisms most likely to improve performance of the METT. We can return to the error data for clues as to the source of error and constraints for future extensions.

Phenomenon 18 states that only a fraction of errors are detectable without reference to the typed copy. Rabbit (1978) reports that omission errors are rarely detected (Table 9) and that the detection of "'de-referencing' of hands" is also extremely low. These two results together indicate that proprioceptive feedback does not give usable information about the force with which a key is hit (too little force produces an omission error, see Phenomenon 21), or about the position of the hand relative to the keyboard. However, usable feedback about the relative position of the fingers as they hit the keys might lead to the possible higher detection rate of substitution, intrusion, and transposition errors. Thus, it seems reasonable that the next level of detail that needs to be added to the METT in order to pick up this phenomenon would introduce feedback about the position of the fingers relative to each other, but not information about the force of a stroke or the position of the hand.

Phenomena 19 through 21 provide a very consistent picture of the most common typing errors. They all indicate that errors occur in physically close proximity to the correct key: substitution errors are mostly adjacent keys, intrusions are often extremely short interkey

intervals because they involve hitting two keys simultaneously, and omissions are mostly long interkey intervals because the omitted key was not hit hard enough. This consistent picture points to the motor processor as the cause of the error. If the perceptual processor misperceived the word, or if the cognitive processor initiated an incorrect letter, then the errors would not primarily occur in the vicinity of the correct letter, but would be distributed across the keyboard. However, if the correct letter were initiated, but incorrectly executed by the motor processor, e.g., with imprecise positioning, the errors would tend to be in the region of the correct key. Therefore, the most profitable direction for expanding the METT to account for errors would be in its model of the motor processor.

Finally, Phenomenon 22 tells us that transposition errors are mostly on alternate hands. Again, a detailed model of the motor processor might shed light on how the sequencing of two signals could become confused. A clue to the mechanism is the timing of the incoming "initiate-letter" signals from the cognitive processor. Since most transposition errors are cross hand, most of the initiation signals from the cognitive processor overlap the motor operator for the previous letter in the motor processor. Within the current METT, this timing is the critical difference between the opportunities for error and the situations where transposition errors rarely occur.

An interesting contemplation about extending the METT's motor model involves the marriage of the METT to Rumelhart and Norman's Activation-Trigger-Schema model of typing (Rumelhart. & Norman, 1982). It might be possible to have the METT's motor model resemble their key press schemata and response system. The METT's cognitive operators, Get-Spelling and Initiate-Letter, might feed activation to the key press schemata with the timing of the onset of activation dictated by the cognitive cycle time. This is pure speculation at this point, but it would be interesting to see if this would preserve the error predictions Rumelhart and Norman obtained. Force and proprioceptive feedback, as well as the details of the hand's physical system suggested by Rumelhart and Norman, might also need to be added to cover the bulk of Salthouse's error phenomena.

Analyzing Sensitivity to Assumptions of the Model

Complete processing models have another advantage for theoretical development; their assumptions are available for sensitivity analysis. For instance, the set of assumptions that make up the METT presented here represent just one set of such assumptions. The METT machine could be changed by modifying one or more assumptions, the crank on each of these new machines could be turned, and the resulting predictions could be examined to see how sensitive the results are to the modifications. What if the same-hand constraint was eliminated so cognitive initiations could proceed without waiting for the motor operators (ASSUMPTION 2b)? Or if, conversely, they had to wait for every motor operator to complete before the next letter could be initiated? Should the size of WM be three words, two, or four (ASSUMPTION 3)? What if the perceptual operator was of different durations for different perceptual chunks (ASSUMPTION 5a)? What if the durations of all types of operators decreased with skill (ASSUMPTION 5c)? These and other questions jump to mind because the model and its assumptions are explicit. Of course, this leads to combinatorial explosion in the work of the theorist beyond the scope of this paper, but at least the rules of the game are clear.

Applying a single model to diverse situations

Returning to Newell's suggestions for psychological research, the third path was to construct a single model of a phenomenon then apply that model in the diverse situations where the phenomenon arises. Although this feature of the METT is not the emphasis of

this paper, the METT is indeed an example of a single model of immediate behavior applied in diverse situations. The same MHP structure of perceptual, cognitive, and motor operators which overlap as information-flow allows that produced the METT, has been used to predict behavior on stimulus-response compatibility tasks ranging from moving a stylus in response to a signal light, to responding verbally with a number in response to the visual presentation of a geometric shape, to typing an abbreviation in response to the visual presentation of a word (John & Newell, 1990). This last task is the one most closely related to typing and provided inspiration for the METT (John & Newell, 1989).

Two other examples of the unity and integration of this single processing model of immediate behavior appear in this analyses of typing phenomena. The analysis of Phenomenon 1 rests on a comparison of an algorithm for a choice reaction-time task and the METT algorithm for typing. This comparison is possible because both algorithms are based on the same type of operators and control structure. More striking is the analysis of Phenomenon 12, where the choice reaction-time task is woven into the typing task. The integration was made with only a few assumptions about which operators take precedence when there was contention for processor resources. This composite model was successful in predicting both the increase in reaction time to the stimulus tone and the relative constancy of interkey interval in the typing task.

Because the METT has been so successful in predicting a broad range of typing phenomena and integrating into a broader model of immediate behavior, it is a contribution to the MHP unified theory of cognition. MHP models of human performance on tasks that require typing can use the METT as a basis for that portion of the task (an example of which will be described below). In addition, the METT served as a guide for an hypothesis of how a more recent and cognitively sophisticated unified theory of cognition, Soar, might model the skill of typing (Newell, 1990). The METT, and the collection of data and defining tasks it modeled, continue to serve as a constraint on the form of Soar models of skilled perceptual-motor behavior.

As a final example of applying a single model to diverse situations, this same framework also simulates the performance of telephone company operators as they use workstations to process toll and assistance calls (e.g., collect, credit card, person-to-person calls) (Gray, et al., 1990, 1991, 1992). These telephone operators talk to customers, type information into a workstation (e.g., the credit-card number), and get information from the workstation (e.g., whether the credit-card is valid). The METT was extended to a more general form that modeled comprehension and generation of speech and the behavior necessary to wait for information that was not always available when it was required for task performance (John, 1990). Models were then used to predict the work times of telephone company operators on two different workstations: the old workstation they were currently using and a new workstation the telephone company was thinking of purchasing. Contrary to expectations, the models predicted and data from an extended field trial confirmed, that asymptotic performance with the new workstation would be <u>slower</u> than with the old workstation. This increase in performance time would have cost the telephone company $2.4 million dollars a year, had they purchased the new workstation (Gray, et al., 1991, 1992). Based on these results, and the explanations for the results provided by the processing models, the telephone company decided not to by the new workstations, but to use models to help them specify better workstations (Atwood, Gray, & John, in press). Models of this type have become a true engineering tool for the evaluation and design of dedicated workstations for performing routine tasks.

That the METT performs so well across different typing tasks is admirable. That it integrates with other forms of the MHP-based model of immediate behavior to allow comparison between tasks and analysis of dual-task performance is impressive. That a

direct descendent of the METT has been used to predict real-world performance and saved an American industry millions of dollars a year, is truly a tribute to Newell's research approach.

ACKNOWLEDGMENTS

This work represents about half of my doctoral dissertation. I wish to thank my advisor, the late Allen Newell, and my committee, John R. Anderson and Lynne M. Reder for their guidance and contributions to this work.

REFERENCES

Atwood, M. E., Gray, W. D., & John, B. E. (in press). Project Ernestine: Analytic and empirical methods applied to a real-world CHI problem. In L. Gugerty & P. Polson (Eds.), Human-Computer Interface Design: Success Cases, Emerging Methods and Real-World Context. San Mateo, CA: Morgan-Kaufmann.

Bovair, S., Kieras, D. E., & Polson, P. G. (1990). The acquisition and performance of text-editing skill: A cognitive complexity analysis. Human-Computer Interaction, 5, 1-48.

Butsch, R. L. C. (1932). Eye movements and the eye-hand span in typewriting. Journal of Educational Psychology, 23, 104-121.

Card, S.K., Moran, T.P., & Newell, A. (1983). The psychology of human-computer interaction. Hillsdale, NJ: Lawrence Erlbaum.

Cheatham, P. G., & White, C. T. (1954). Temporal numerosity: III. Auditory perception of number. Journal of Experimental Psychology, 47, 425-428.

Cooper, W. E. (1983). Cognitive aspects of skilled typewriting. New York: Springer-Verlag.

Gentner, D. R. (1983). The acquisition of typewriting skill. Acta Psychologica, 54, 233-248.

Gray, W. D., John, B. E., & Atwood, M. E. (1991). Project Ernestine: A validation of GOMS for prediction and explanation of real-world task performance. Manuscript submitted for publication.

Gray, W. D., John, B. E., & Atwood, M. E. (1992). The precis of Project Ernestine or an overview of a validation of GOMS. In P. Bauersfeld, J. Bennett & G. Lynch (Eds.) Proceedings of CHI 1992 (pp. 307-312). Monterey, CA: ACM.

Gray, W. D., John, B. E., Stuart, R., Lawrence, D., & Atwood. M. E. (1990). GOMS meets the phone company: Analytic modelling applied to real-world problems. In D. Diaper, D. Gilmore, G. Cockton, & B. Shackel (Eds.), Human-Computer Interaction INTERACT '90 (pp. 29-34). Cambridge, U.K.: North-Holland Elsevier Science Publishers.

Henderson, L. (1982). Orthography and word recognition in reading. London: Academic Press.

Hershman, R. L., & Hillix, W. A. (1965). Data processing in typing: Typing as a function of kind of material and amount exposed. Human Factors, 7, 483-292.

John, B. E. (1988). Contributions to engineering models of human-computer interaction. Unpublished doctoral dissertation, Carnegie Mellon University, Pittsburgh. (Also appeared as ONR Technical Report A.D. No. A195590).

John, B. E. (1990). Extensions of GOMS analyses to expert performance requiring perception of dynamic visual and auditory information. Proceedings of CHI 1990 (pp. 107-115). Seattle, WA: ACM.

John, B. E., & Newell, A. (1987). Predicting the time to recall computer command abbreviations. In J. M. Carroll & P. P. Tanner (Eds.), Proceedings of CHI+GI 1987 (pp. 33-40). Toronto, Ontario, Canada: ACM.

John, B. E., & Newell, A. (1989). Cumulating the science of HCI: From S-R compatibility to transcription typing. In K. Bice & C. Lewis (Eds.), Proceedings of CHI 1989 (pp. 109-114). Austin, TX: ACM.

John, B. E., & Newell, A. (1990). Stimulus-response compatibility in a unified theory of cognition. In R. W. Proctor & T. G. Reeve (Eds.), Stimulus-response compatibility: An integrated approach (pp. 427-479) Amsterdam: North-Holland Elsevier Science Publishers.

John, B. E., Rosenbloom, P. S., & Newell, A. (1985). A theory of stimulus-response compatibility applied to human-computer interaction. In L. Borman & B. Curtis (Eds.), Proceedings of CHI 1985 (pp. 213-219). San Francisco, CA: ACM.

Kieras, D. E., & Polson, P. G. (1985). An approach to the formal analysis of user complexity. International Journal of Man-Machine Studies, 22, 365-394.

Larochelle, S. (1983). A comparison of skilled and novice performance in discontinuous typing. In W. E. Cooper (Ed.), Cognitive aspects of skilled typewriting (pp. 67-94). New York: Springer-Verlag.

Lerch, F. J., Mantei, M. M., & Olson, J. R. (1989). Skilled financial planning: The cost of translating ideas into action. In K. Bice & C. Lewis (Eds.), Proceedings of CHI 1989 (pp. 121-126). Austin, TX: ACM.

Logan, G. D. (1982). On the ability to inhibit complex movements: A stop-signal study of typewriting. Journal of Experimental Psychology: Human Perception and Performance, 8(6), 778-792.

Logan, G. D. (1983). Time, information, and the various spans in typewriting. In W. E. Cooper (Ed.), Cognitive aspects of skilled typewriting (pp. 197-224). New York: Springer-Verlag.

Luce, R. D. (1986). Response times: Their role in inferring elementary mental organization. New York: Oxford University Press.

Newell, A. (1973). You can't play 20 questions with nature and win: Projective comments on the papers of this symposium. In W. G. Chase (Ed.), Visual Information Processing (pp. 283-308). New York: Academic Press.

Newell, A. (1990). Unified Theories of Cognition. Cambridge, MA: Harvard University Press.

Newell, A., & Card, S. K. (1985). The prospects for psychological science in human-computer interaction. Human-Computer Interaction, 1, 209-242.

Newell, A., & Card, S. K. (1986). Straightening out softening up: Response to Carroll and Campbell. Human Computer Interaction, 2, 251-267.

Polson, P. G., & Kieras, D. E. (1985). A quantitative model of the learning and performance of text editing knowledge. In L. Borman & B. Curtis (Eds.), Proceedings of CHI 1985 (pp. 207-212). San Francisco, CA: ACM.

Ostry, D. J. (1983). Determinants of interkey times in typing. In W. E. Cooper (Ed.), Cognitive aspects of skilled typewriting (pp. 225-246). New York: Springer-Verlag.

Rabbit, P. (1978). Detection of errors by skilled typists. Ergonomics, 21(11), 945-958.

Rothkopf, E. Z. (1980). Copying span as a measure of the information burden in written language. Journal of Verbal Learning and Verbal Behavior, 19, 562-572.

Rumelhart, D. E., & Norman, D. A. (1982). Simulating a skilled typist: A study of skilled cognitive-motor performance. Cognitive Science, 6, 1-36

Salthouse, T. A. (1984a). Effects of age and skill in typing. Journal of Experimental Psychology: General, 113(3), 345-371.

Salthouse, T. A. (1984b). The skill of typing. Scientific American, 250, 128-135.

Salthouse, T. A. (1985). Anticipatory processes in transcription typing. Journal of Applied Psychology, 70, 264-271.

Salthouse, T. A. (1986). Perceptual, cognitive, and motoric aspects of transcription typing. Psychological Bulletin, 99(3), 303-319.

Salthouse, T. A., & Saults, J. S. (1987). Multiple spans in transcription typing. Journal of Applied Psychology, 72(2), 187-196.

Shaffer, L. H. (1973). Latency mechanisms in transcription. In S. Kornblum (Ed.), Attention and performance, IV (pp. 435-446). New York: Academic Press.

Shaffer, L. H. (1975). Control processes in typing. Quarterly Journal of Experimental Psychology, 27, 419-432.

Shaffer, L. H. (1976). Intention and performance. Psychological Review, 83, 375-393.

Shaffer, L. H., & French, A. (1971). Coding factors in transcription. Quarterly Journal of Experimental Psychology, 23, 268-274.

Shaffer, L. H., & Hardwick, J. (1970). The basis of transcription skill. Journal of Experimental Psychology, 84, 424-440.

Smelcer, J. B. (1989). Understanding user errors in database query. Unpublished doctoral dissertation, University of Michigan, Ann Arbor.

Sternberg, S., Knoll, R. L., & Wright, C. E. (1978). Experiments on temporal aspects of keyboard entry. In J. P. Duncanson (Ed.), Getting it together: Research and applications in human factors, (pp. 28-50). Santa Monica, CA: Human Factors Society.

Thomas, E. A., & Jones, R. G. (1970). A model for subjective grouping in typewriting. Quarterly journal of Experimental Psychology, 22, 353-367.

Webster's New Collegiate Dictionary. (1979). Springfield, Mass.: G. & C. Merriam Company.

West, L. J. and Sabban, Y. (1982). Hierarchy of stroking habits at the typewriter. Journal of Applied Psychology, 67, 370-376.

Ziegler, J. E., Hoppe, H. U., & Fahnrich, K. P. (1986) Learning and transfer for text and graphics with a direct manipulation interface. In M. Mantei and P. Orbeton (Eds.), Proceedings of CHI'86 (pp. 72-77). Boston, MA: ACM.

Tables and Figures

Table 1. Perceptual, cognitive and motor parameter definitions and estimated durations.

Parameter	Definition	Duration
Perceptual Operator	Reading a word of about 6 letters and encoding it into an ordered list of letters	340 msec
Cognitive Operator	A cognitive processor cycle time	50 msec
Motor Operator	Typing a character on an alphanumeric keyboard at a rate of about 30 gwpm	230 msec

Table 2. Effective typing speeds with each of the strategies.

	Effective Typing Speed (wpm)				
	-------------Strategies--------------			Straight	Weighted*
	N	3-p-l	3-l	average	average
EP	98	-	-	98	98
LC	99	84	45	76	65
LJ	71	57	45	58	51

*"weighted" means averaged across strategies 3-p-l and 3-l only for LC and LJ.

Table 3. Observed and predicted percentage change between typing conditions.

	Observed % Change (sd)		Predicted Straight Average [%err]		Predicted Weighted* Average [%err]	
EP vs. LJ	94.4	(29.4)	69.0	[26.9]	92.2	[2.3]
EP vs. LC	61.9	(22.1)	28.9	[53.3]	50.8	[17.9]
LC vs. LJ	20.5	(10.9)	31.0	[51.2]	27.5	[34.1]
Average absolute % error				43.8		18.1

*weighted" means only averaged across strategies 3-p-l and 3-l for LC and LJ.

Table 4. Comparison of time predictions made by the METT and Rumelhart & Norman's simulation.

			-------MHP (Fitts's Law)------			-----Rumelhart & Norman-----		
Keys	Tot Time Obs (msec)	Dist. Moved (inches)	Est. Move Time (msec)	Total Time Pred (msec)	%err	Simulation Units	Total Time Pred (msec)	%err
U.	165	0.00	0	166	-0.6	5.9	159	3.6
d-e	201	0.79	34	195	3.0	9.5	192	4.5
c-e	215	1.58	55	221	-2.8	12.7	221	-2.8
r-e	145	0.40	14	146	-0.7	7.1	170	-17.2
t-e	159	0.63	28	157	1.3	7.0	169	-6.3
f-e	168	0.79	35	162	3.6	7.3	171	-1.8
g-e	178	0.98	41	174	2.2	7.2	171	3.9
v-e	178	1.19	47	180	-1.1	7.5	173	2.8
b-e	195	1.35	50	183	6.2	8.1	179	8.2
			Average absolute % error = 2.4			Average absolute p% error = 5.7		

Table 5. Results of typing and reaction-time tasks, alone and concurrent.

	alone	concurrent
typing interkey interval (sd), in msec.	181 (64)	185 (62)
reaction time (sd), in msec.	269 (49)	431 (85)

Table 6. Copy span predictions for an 80 gwpm typist.

Character Typed	Available to Type	Copy span
O	ne_reason_	10
n	e_reason_is_	9
e	_reason is_	8
_	reason_is_	10
r	eason_is_quite_	9
e	ason_is_quite_	14
a	son_is_quite_	13
s	on_is_quite_	12
o	n_is_quite_	11
n	_is_quite_	10
_	is_quite_	9
i	s_quite_	8
s_	_quite_obvious;_	16

Table 7. Stopping span predictions for three different tasks.

Experiment Number	2	1	3
Signal Contingency	event	time	event
Context	single word	single word	sentence
Observed Stopping Span	1.41	1.57	2.16
Predicted Stopping Span	1.55	1.76	2.08
Percent Error	-9.9%	-12.1%	3.7%

Table 8. Interkeystroke interval (msec) for different size preview windows for a 120 gwpm typist.

Preview Characters	Q1	Q2
unlimited	78	96
AA.	78	96
11	83	103
9	96	119
7	104	128
5	138	170
3	200	247
1	369	456

Table 9. Errors made and detected in Rabbit's experiment (1978).

	Errors Made	Errors Corrected	Percentage Corrections
Grand Total of all Errors	7089	4447	62.73%
"Compound Errors" involving two or more responses	3403	1965	57.74%
Single letter mistypes	3591	2478	69.00%
Omission Errors	95	4	4.21%

Table 10. Interkeystroke interval (msec) for different size preview windows for 60, 90 and 120 gwpm typists.

	60 gwpm		90 gwpm		120 gwpm	
Preview Character	Q1	Q2	Q1	Q2	Q1	Q2
unlimited	158	195	109	135	78	96
PP.	158	195	109	135	78	96
11	159	196	112	138	83	103
9	170	210	124	153	96	119
7	177	218	133	164	104	128
5	190	234	156	193	138	170
3	247	305	215	266	200	247
1	450	556	402	496	369	456

Table 11. Phenomena accounted for by the METT.

Type of Prediction	Type of Phenomenon		Number
Parametric	Parametric		3
Quantitative	Quantitative		7
Qualitative	Qualitative		6
Qualitative	Quantitative		5
Not covered			8
Of all 29 phenomena:			
Total as good as the data		16/29	55%
Total accounted for at least qualitatively		21/29	72%
Of the 21 phenomena covered by the METT:			
Total as good as the data		16/21	76%
Total accounted for at least qualitatively		21/21	100%

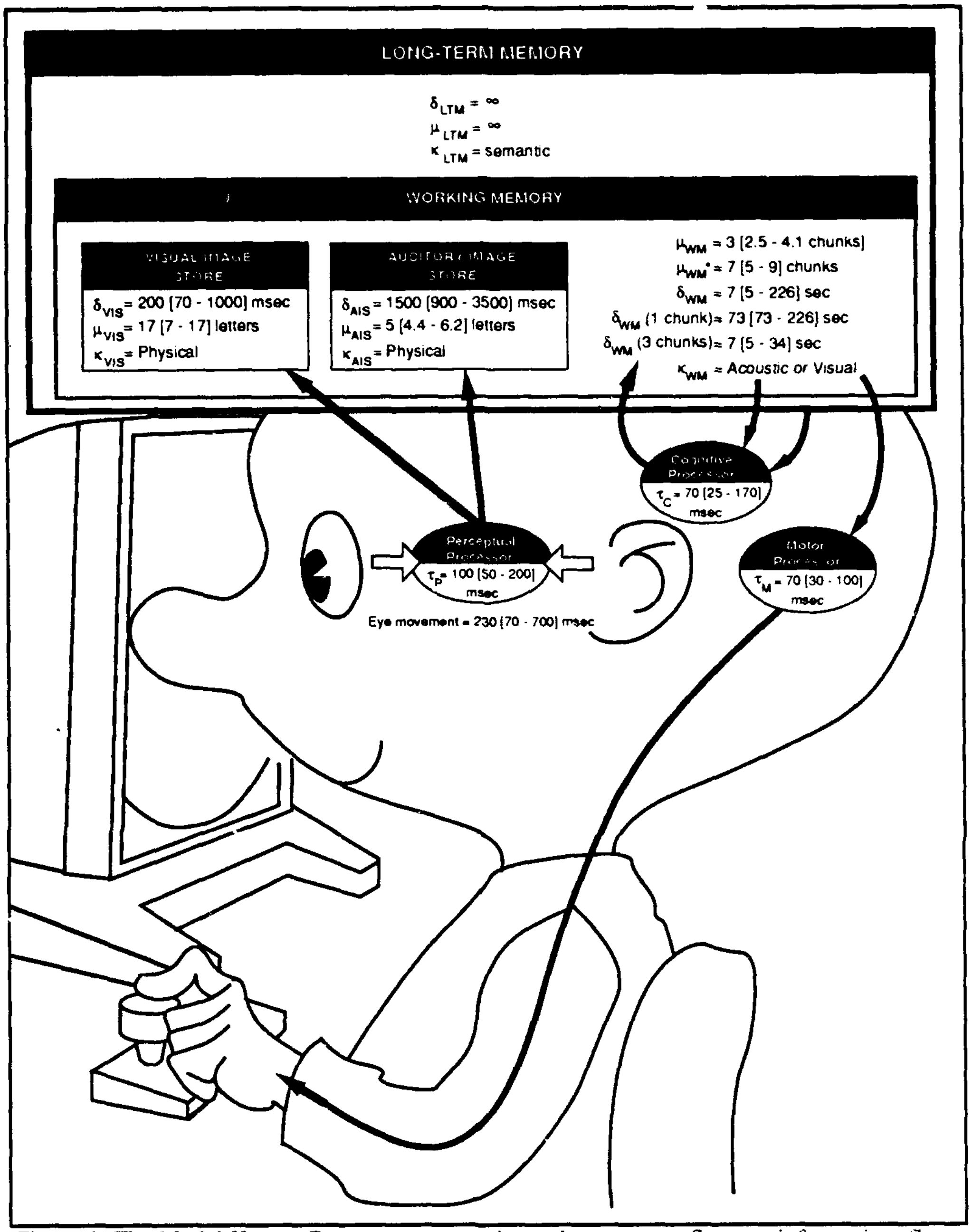

Figure 1. The Model Human Processor - memories and processors. Sensory information, flows into Working Memory through the Perceptual Processor. Working Memory consists of activated chunks in Long-Term Memory. The basic principle of operation of the Model Human Processor is the Recognize-Act Cycle of the Cognitive Processor: On each cycle of the Cognitive Processor, the contents of Working Memory initiate actions associatively linked to them in Long-Tem Memory; these actions in turn modify the contents of Working Memory. The Motor Processor is set in motion through activation of chunks in Working Memory. Predictions are made using a set of Principles of Operations: (PO) the Recognize-Act Cycle of the Cognitive Processor, (P1) the Variable Perceptual Processor Rate Principle, (P2) The Encoding Specificity Principle, (P3) The Discrimination Principle, (P4) The Variable Cognitive Processor Rate Principle, (P5) Fitt's Law, (P6) the Power Law of Practice, (P7) The Psychological Uncertainty Prinicple, (P8) The Rationality Principle, and (P9) The Problem Space Principle. (Adapted from Card, Moran & Newell (1983, Figures 2.1, 2.2) with permission.)

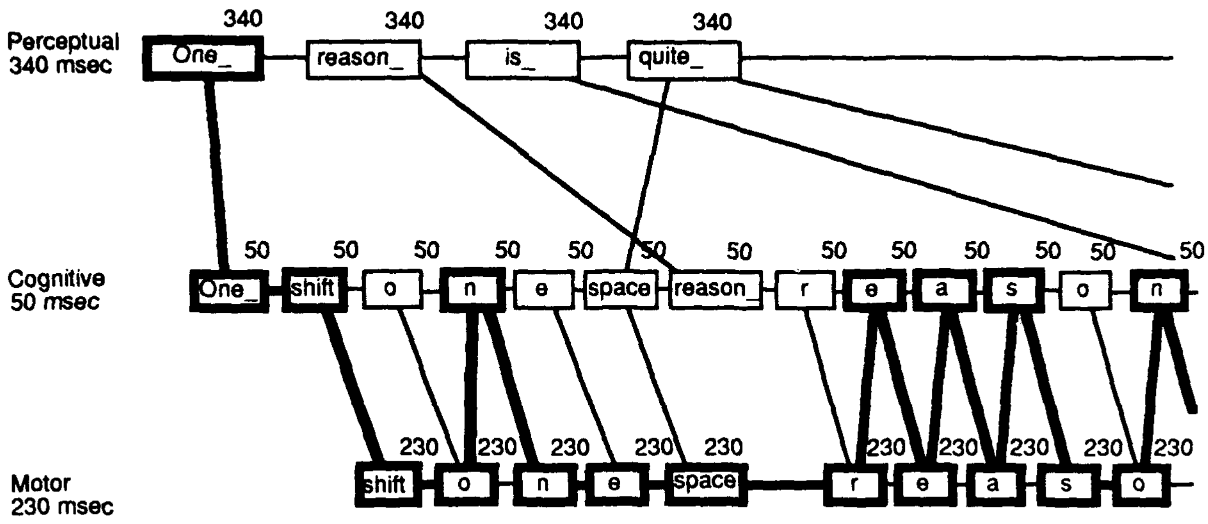

Figure 2. A schedule chart and critical path for the example sentence for a 60 gwpm typist with an initial motor operator estimate of 230 msec.

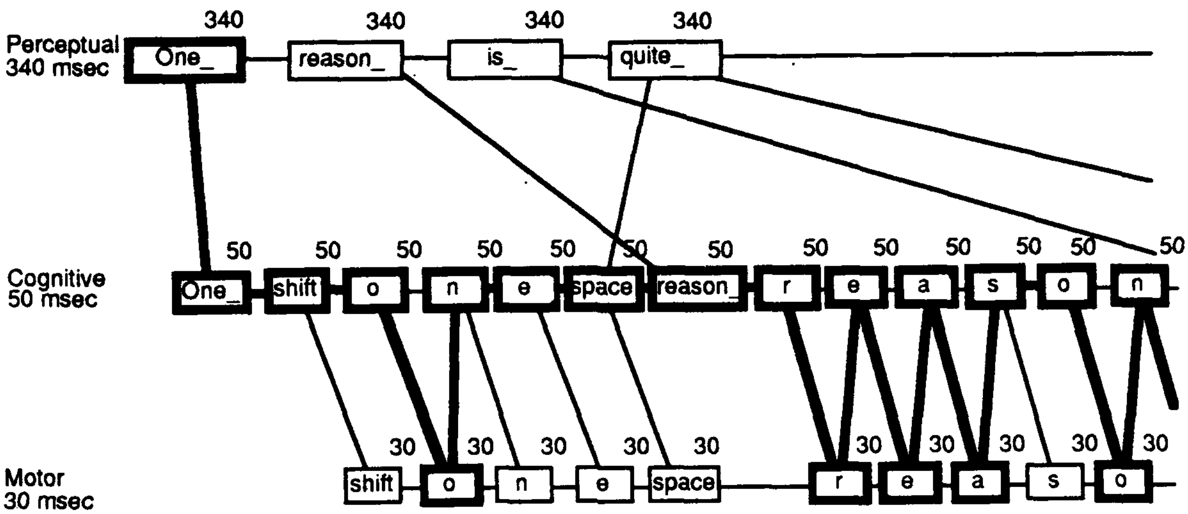

Figure 3. A schedule chart and critical path for the example sentence for a 160 gwpm typist.

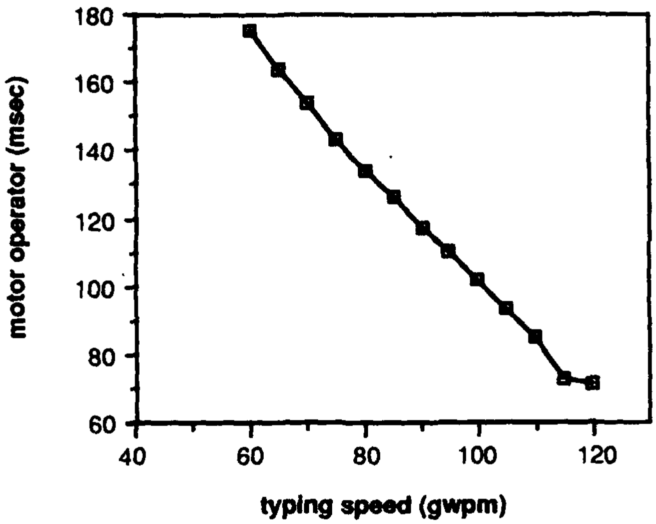

Figure 4. Typing speed vs. motor operator estimate.

EP	**(Easy prose)**	I have your letter in which you ask about the prices...
LC	**(Letter comb)**	I veha uryo terlet ni chwhi ouy ska outab eth espric...
LJ	**(Letter jumble)**	I evah uoyr rtleet ni hcihw oyu ska auobt teh rpcsei...
EW	**(Easy words)**	Letter the I of about next ask in have month your which...
RW	**(Rare words)**	Tycoon alp a si gumbo jamb boa em plop joist ouch piker...

Figure 5. West and Sabban's (1982) five types of materials.

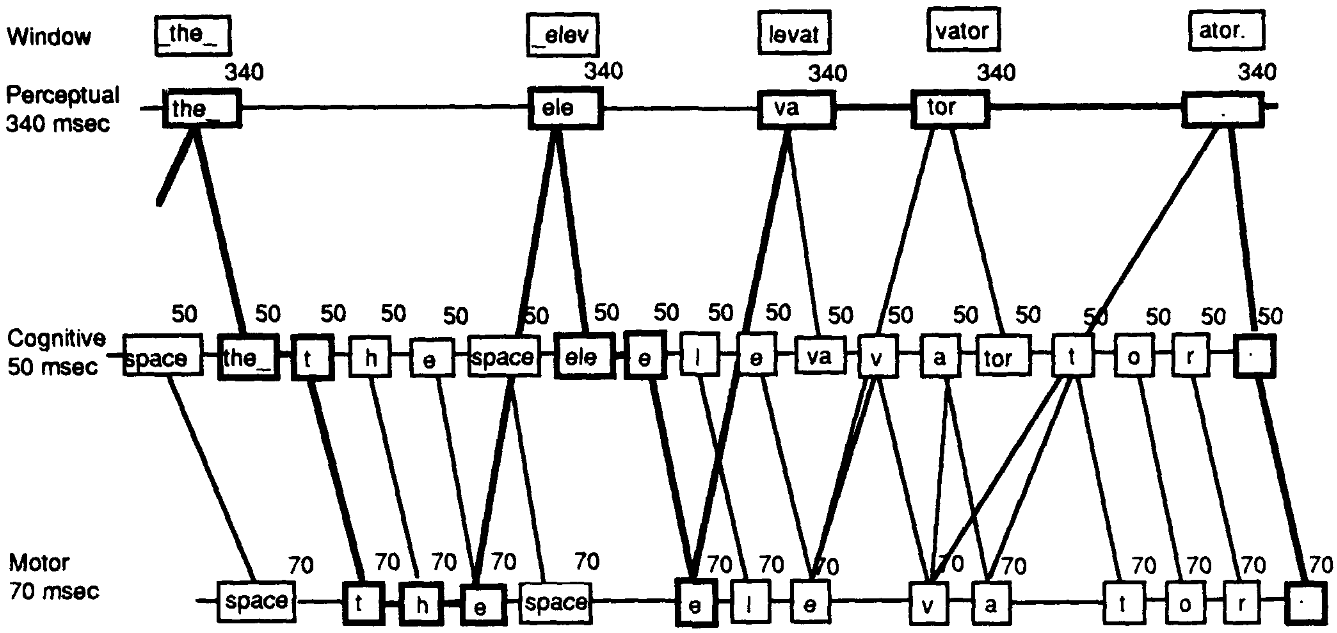

Figure 6. Schedule chart for "the elevator" in a five-character preview

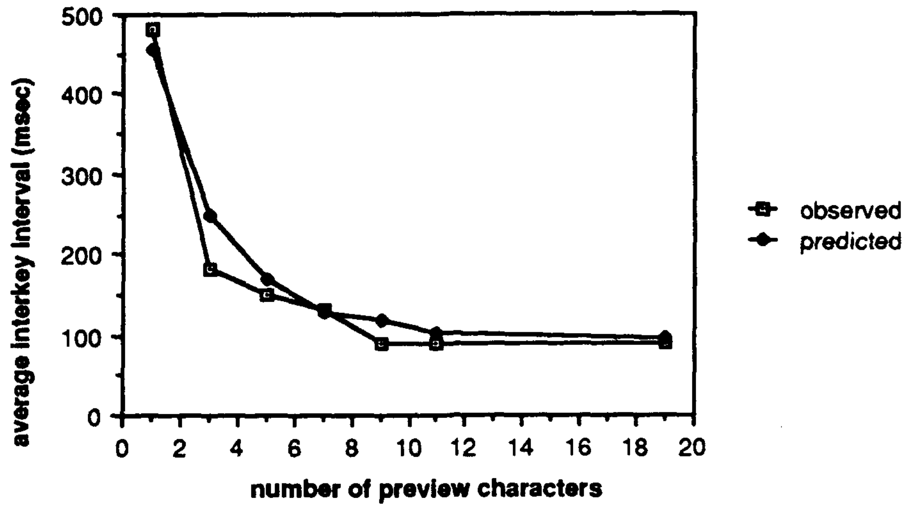

Figure 7. Restricted preview predicted and observed results.

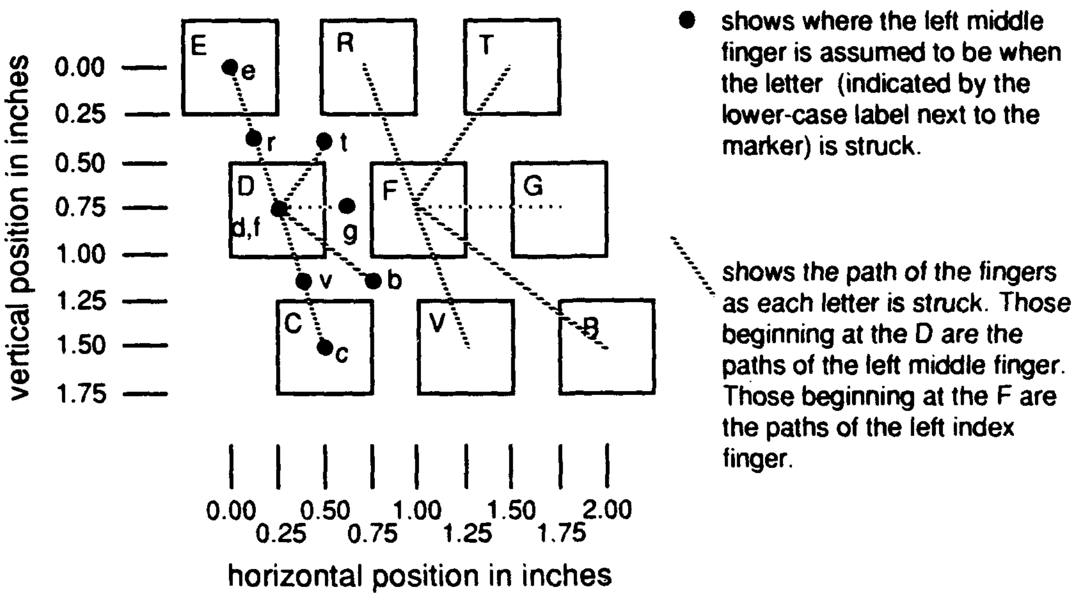

Figure 8. Keyboard and position of the left middle finger when each key is hit.

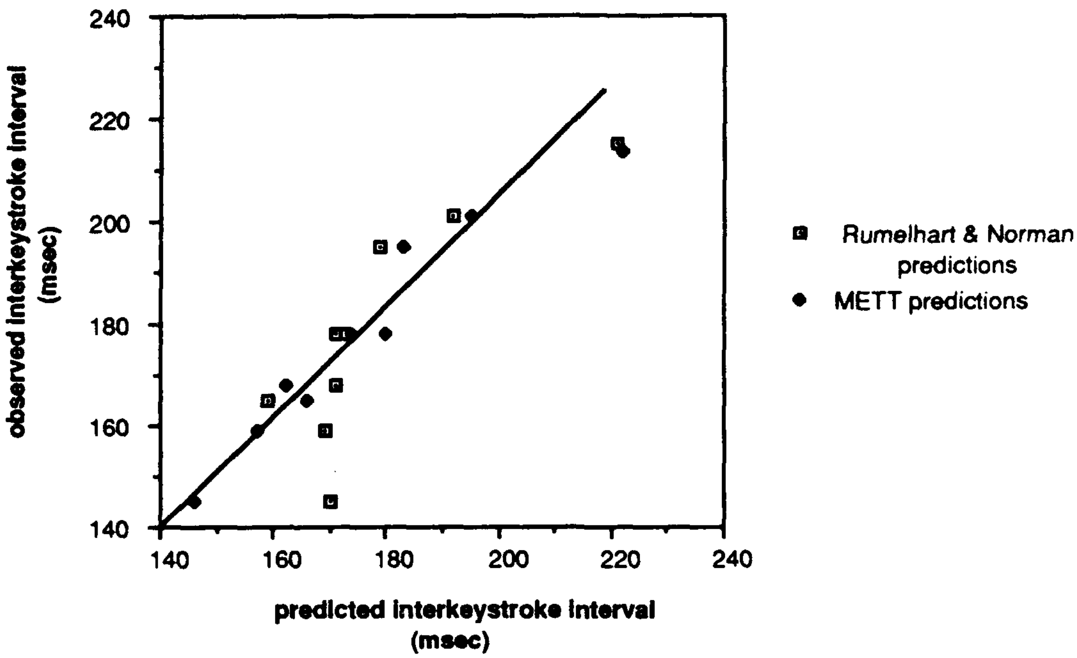

Figure 9. Predictions of movement time made by Rumelhart & Norman's model and by the METT.

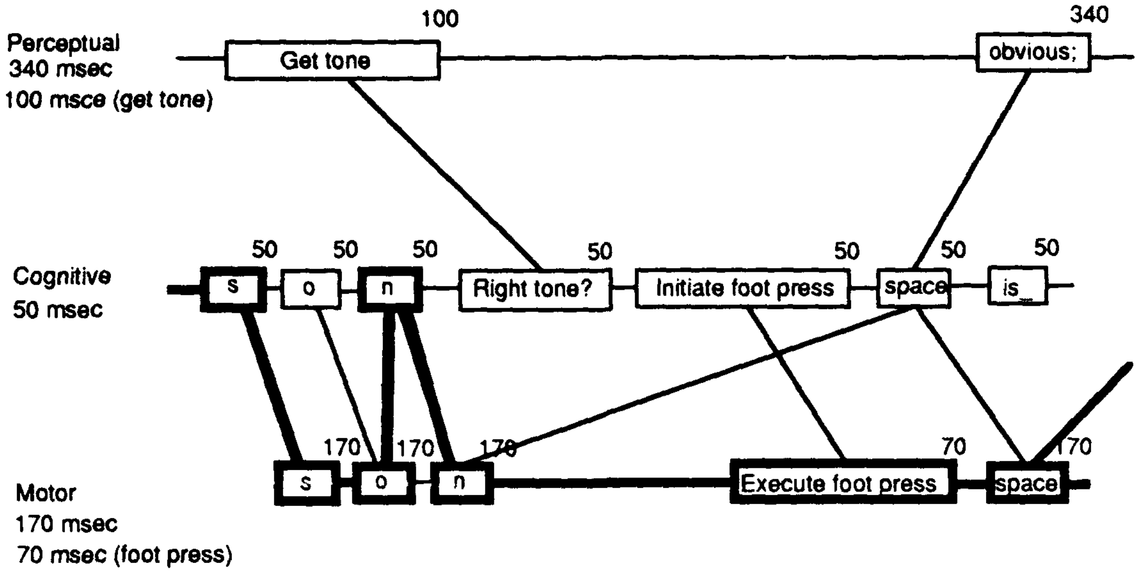

Figure 10. Schedule chart for the concurrent typing and reaction-time tasks.

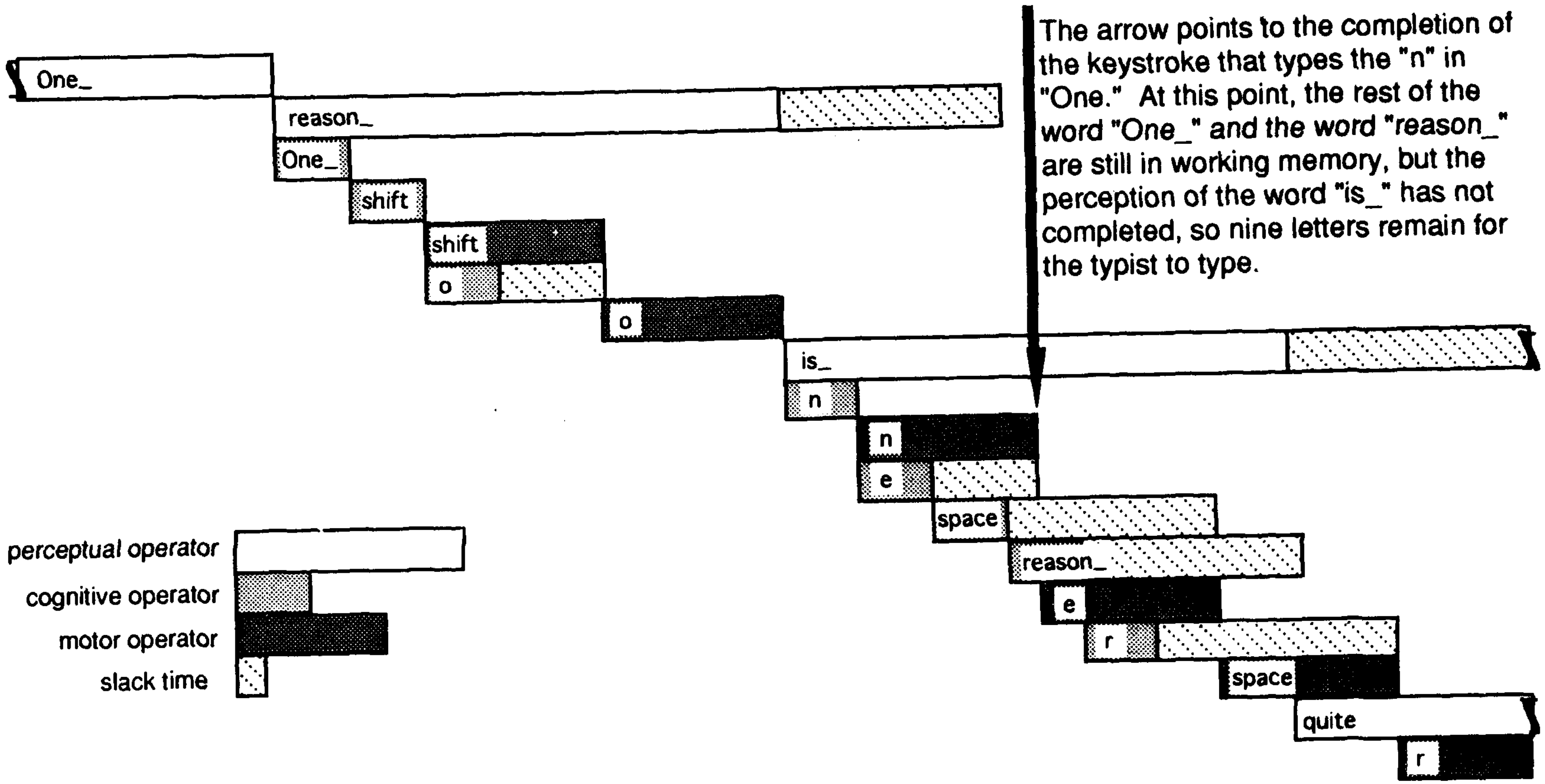

Figure 11. Portion of a task timeline for an 80 gwpm typist showing the copy span analysis.

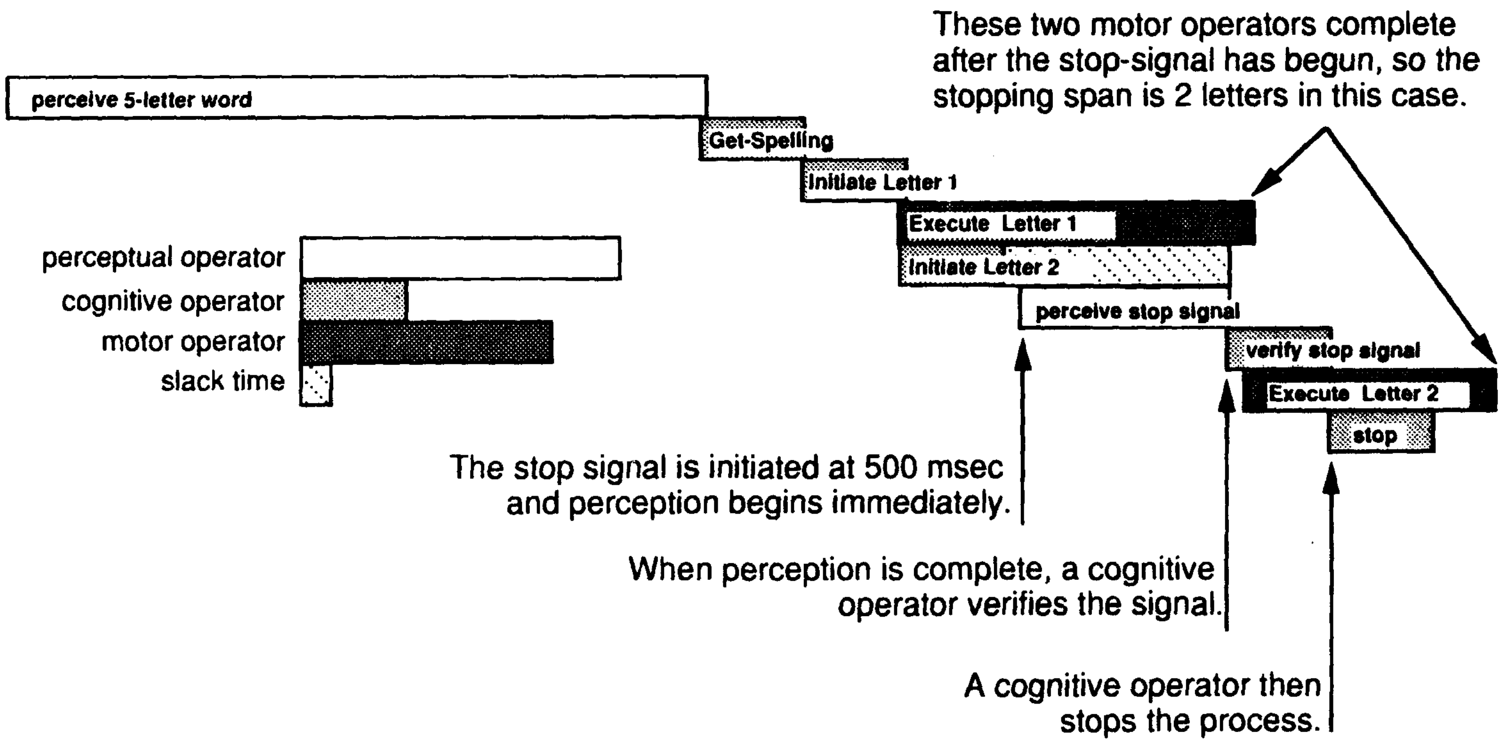

Figure 12. Timeline of a 60 gwpm typist, a stop-signal at 500 msec, and a word whose first two letters are on opposite hands and whose third letter is on the same hand as the second.

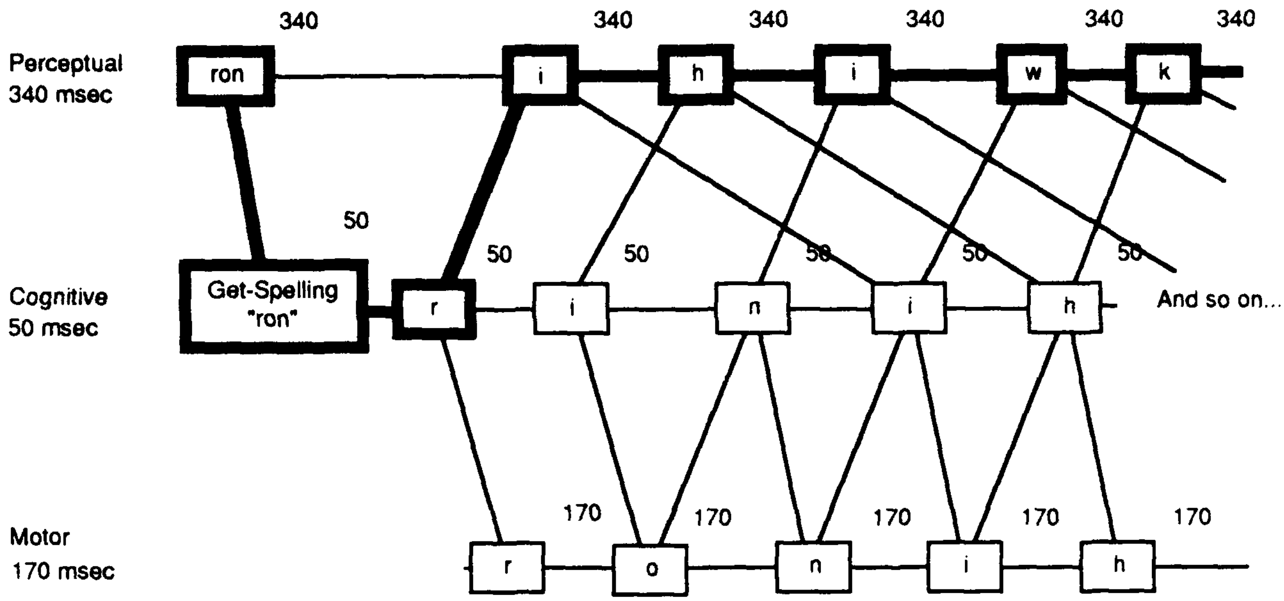

Figure 13. Portion of a schedule chart for typing random letters.

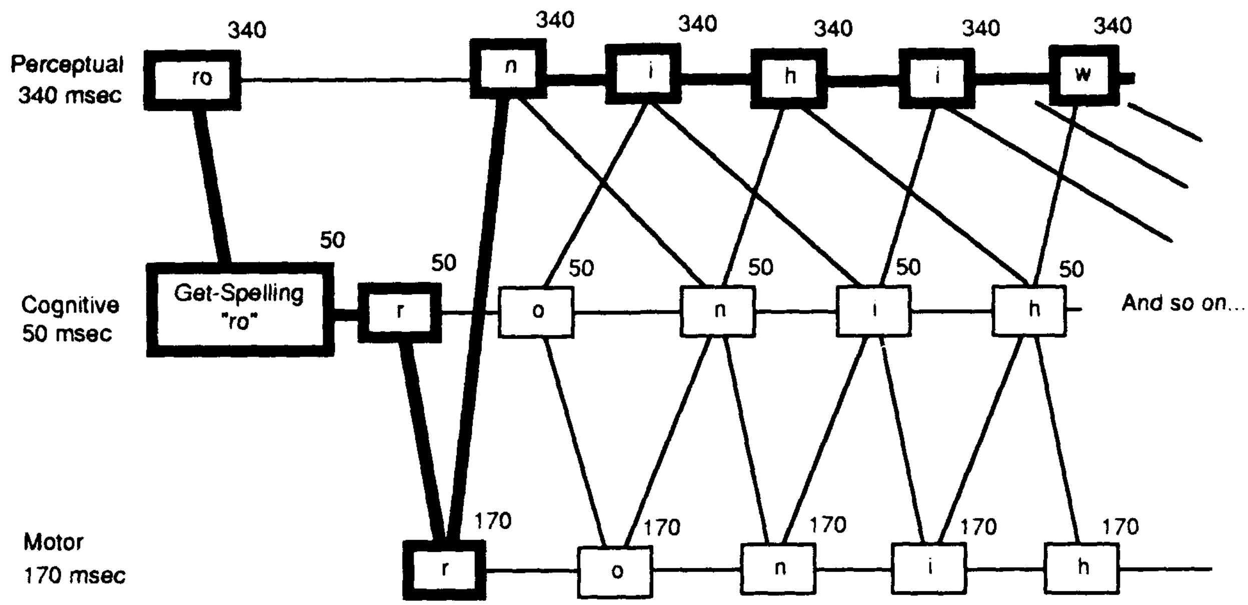

Figure 14. Portion of a schedule chart for typing random letters with a 2-letter preview window.

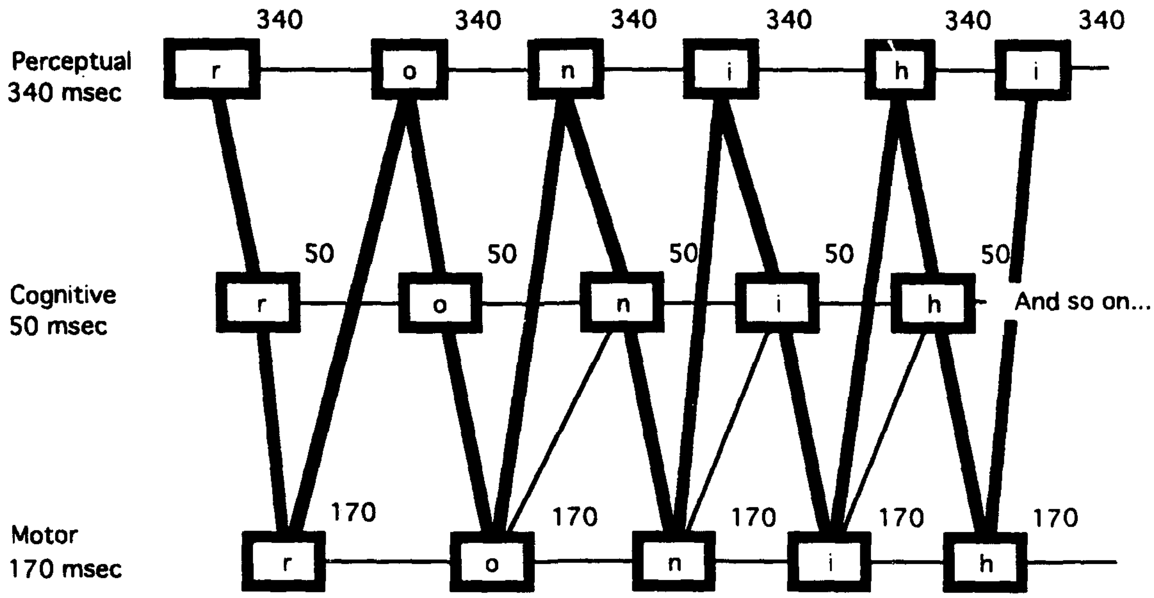

Figure 15. Portion of a schedule chart for typing random letters with a 1-letter preview window.

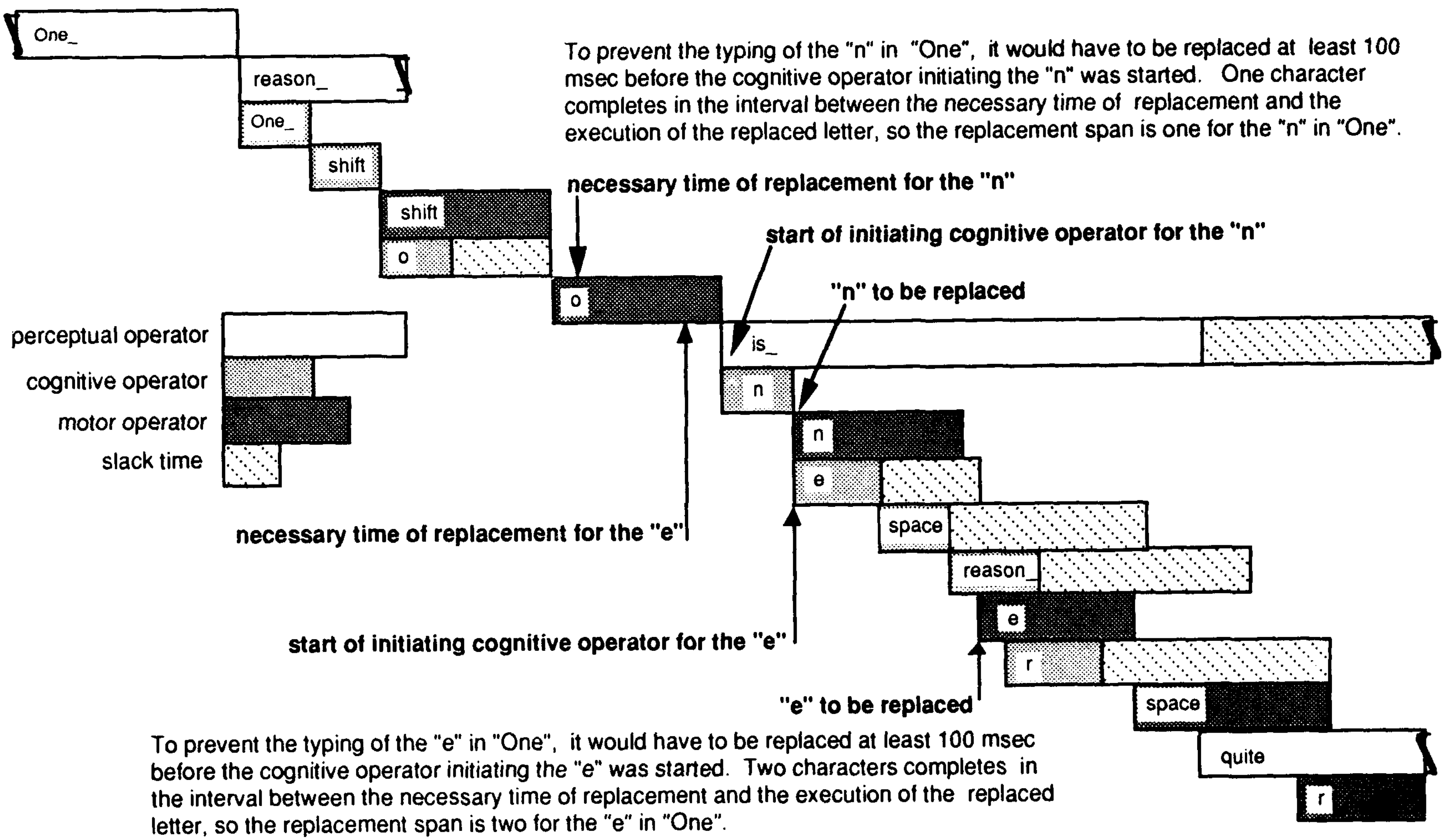

Figure 16. Portion of a task timeline for an 60 gwpm typist showing the replacement span analysis.

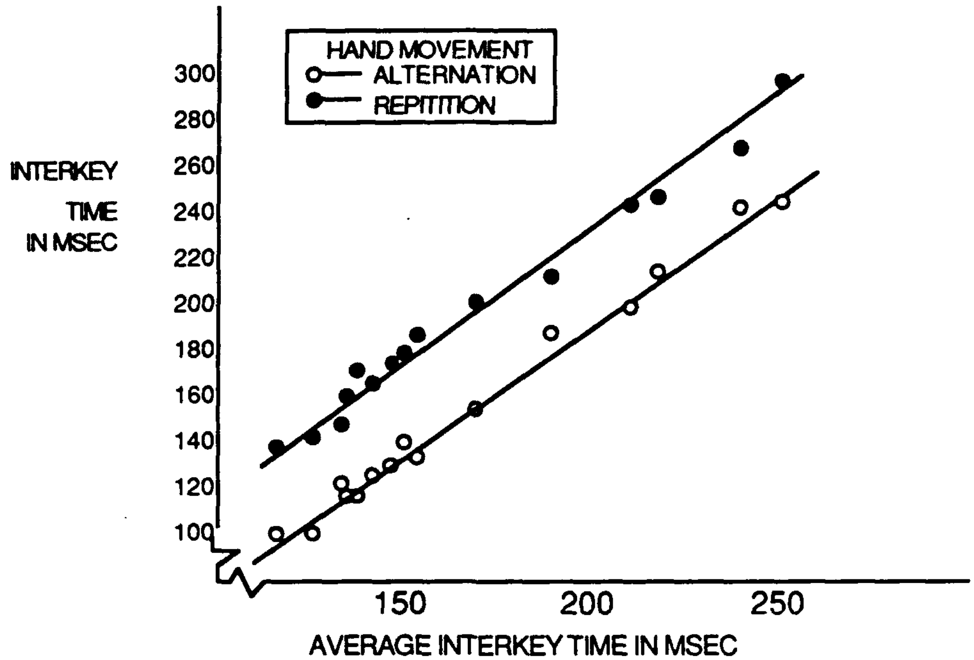

Figure 17. Average interkey times as a function of typing speed (Ostry, 1983).

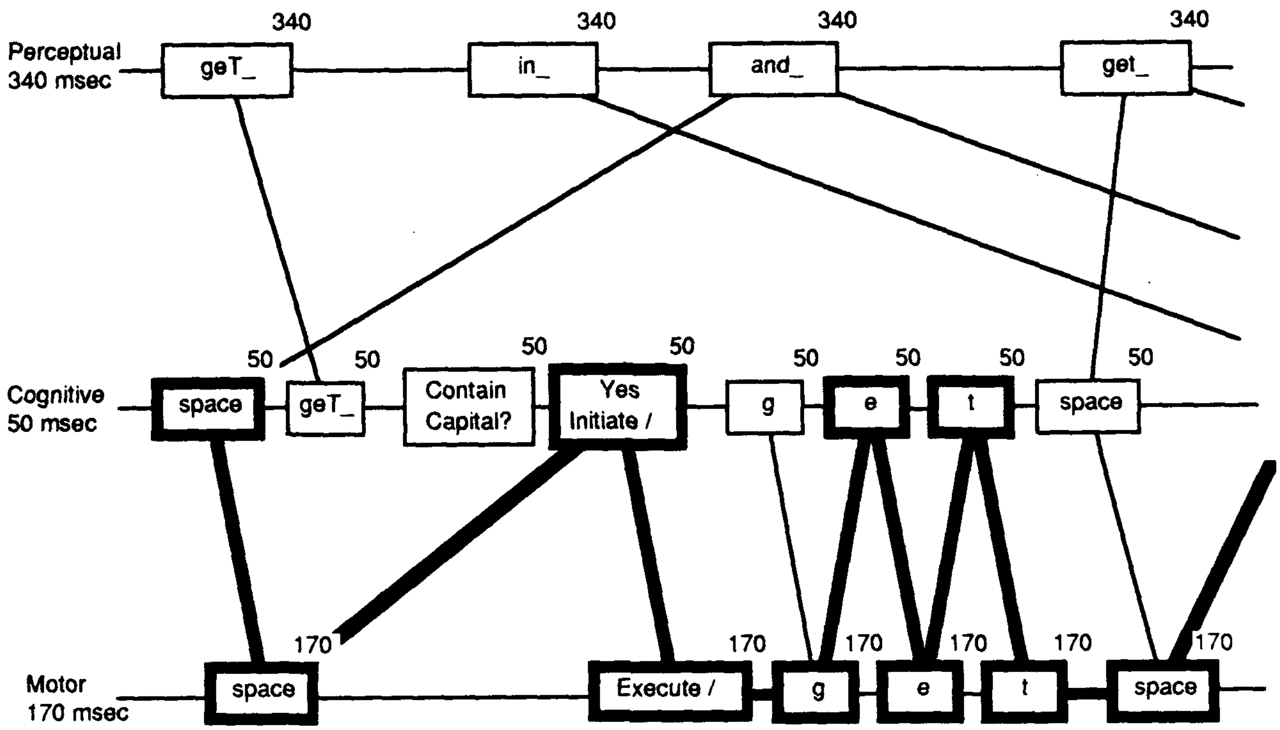

Figure 18. Schedule chart of the spelling algorithm for the detection span task.

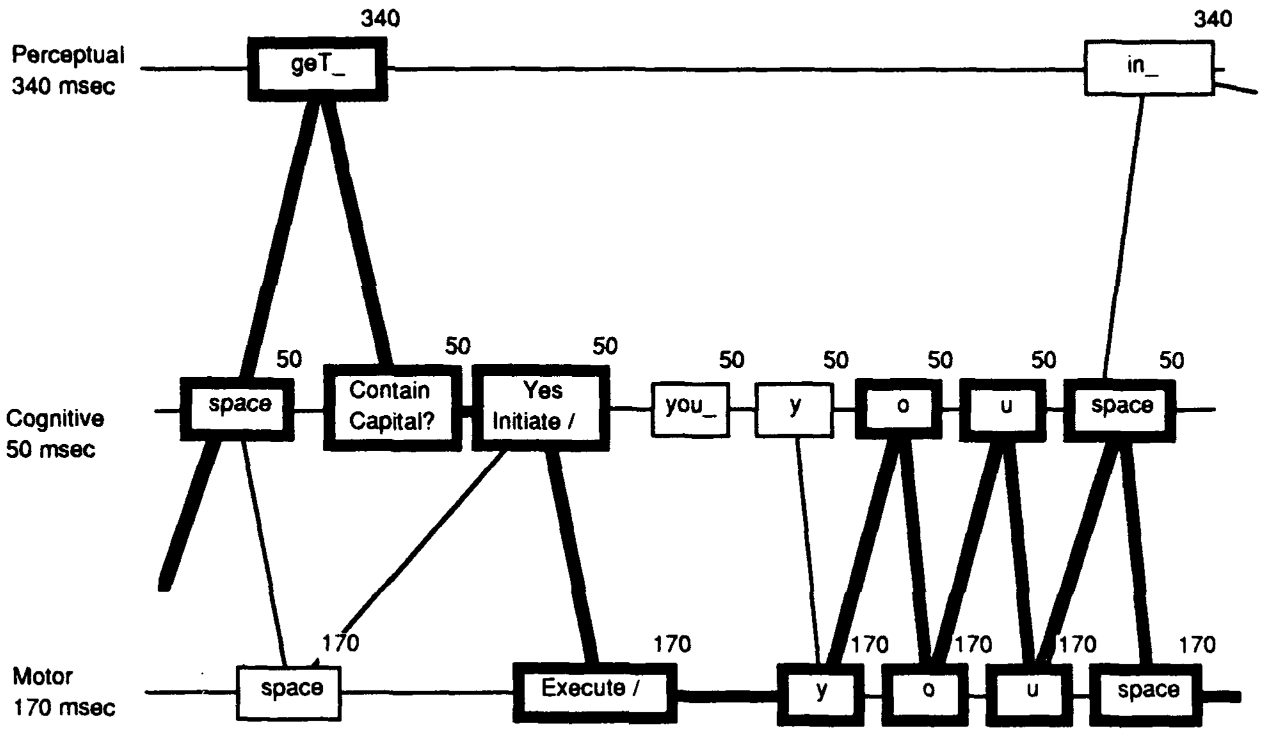

Figure 19. Schedule chart of the perception-wait algorithm for the detection span task.

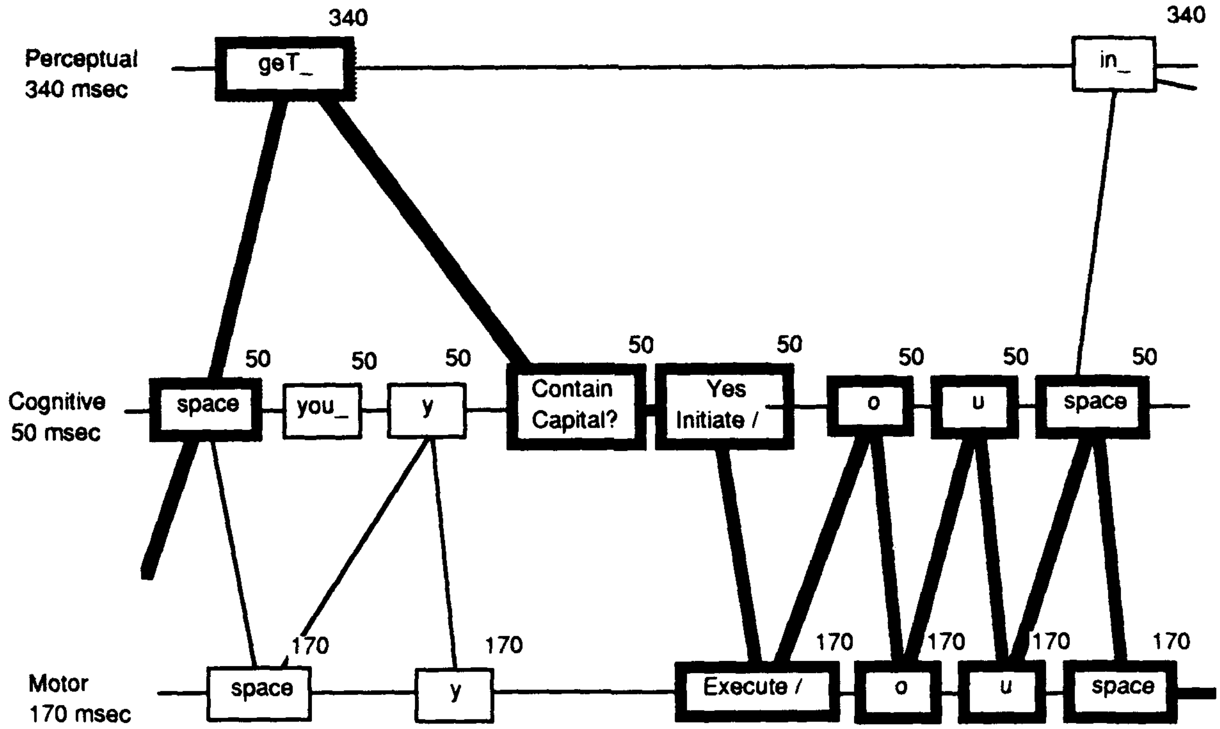

Figure 20. Schedule chart of the perception-parallel algorithm for the detection span task.

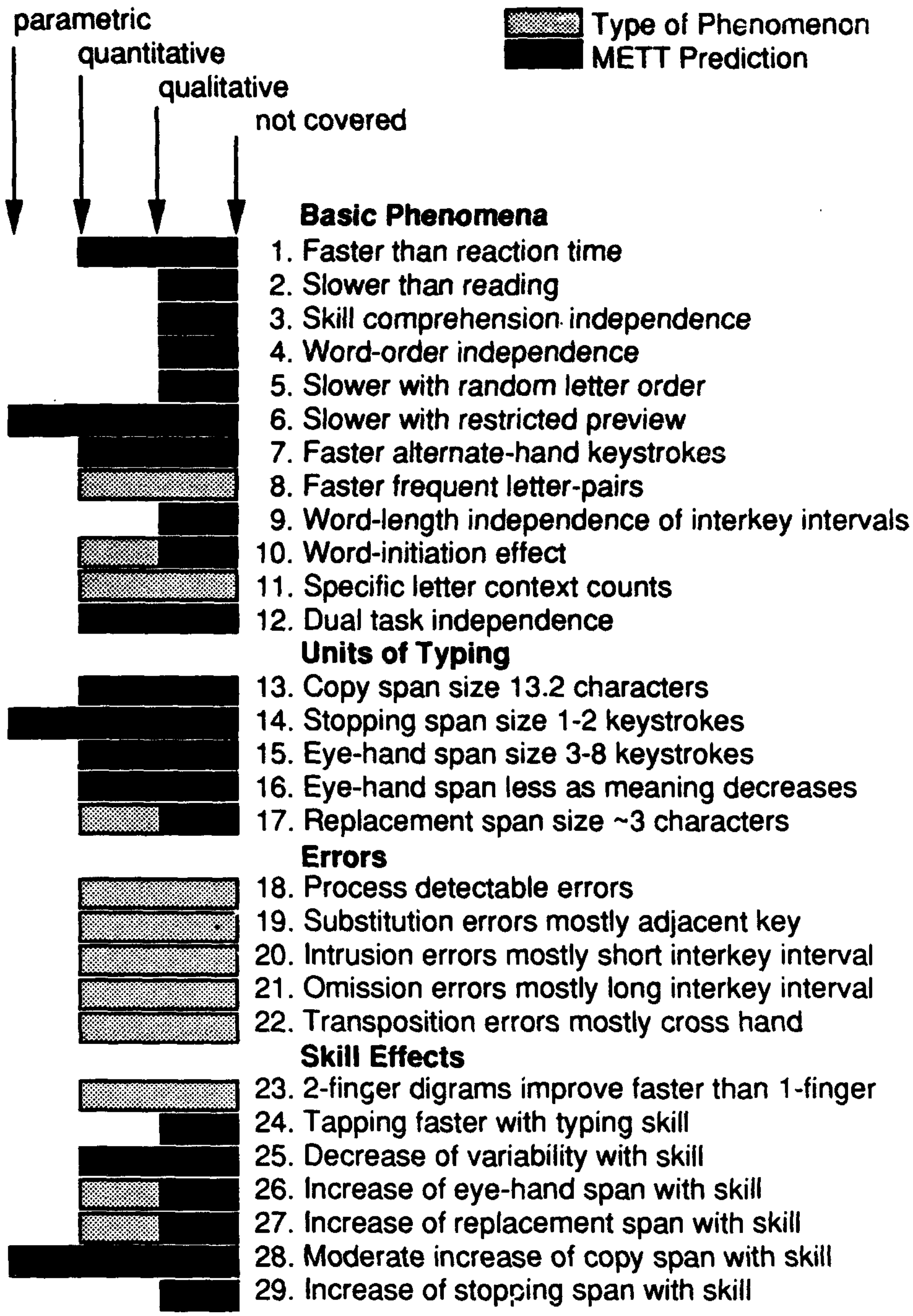

Figure 21. Summary of the METT account of the Salthouse 29.